DOCUMENT RESUME

ED 160 238 PS 010 166

AUTHOR	Granville, Arthur C.; And Others
TITLE	A Process Evaluation of Project Developmental Continuity. Interim Report IV, Volume I: Pilot Year Impact Study--Instrument Characteristics and Attrition Trends.
INSTITUTION	High/Scope Educational Research Foundation, Ypsilanti, Mich.
SPONS AGENCY	Office of Child Development (DHEW), Washington, D.C. Early Childhood Research and Evaluation Branch.
PUB DATE	Aug 76
CONTRACT	HEW-105-75-1114
NOTE	280p.; For related documents, see ED 144 715, PS 010 163-165, and PS 010 167-176; This series includes all the public reports generated by this study
EDRS PRICE	MF-$0.83 HC-$15.39 Plus Postage.
DESCRIPTORS	Attrition (Research Studies); Comparative Analysis; Data Collection; *Demonstration Programs; *Early Childhood Education; Educational Research; *Evaluation Methods; Formative Evaluation; Matched Groups; *Measurement Instruments; *Program Evaluation; Research Methodology; Sampling; Statistical Analysis; Tables (Data); Test Reliability; Test Validity
IDENTIFIERS	*Developmental Continuity; *Project Developmental Continuity; Project Head Start

ABSTRACT

This interim report of a pilot year impact study on evaluation methodology is part of a series of documents on the evaluation of Project Developmental Continuity, a Head Start demonstration program aimed at promoting educational and developmental continuity between children's Head Start and primary school experiences. This report deals with instrument suitability, group comparability, and sample attrition. Chapter I gives an overview of PDC and describes the purpose of the PDC evaluation; Chapter II describes data collection and data analysis procedures; Chapter III presents findings on the general utility of the test instruments, results of comparability checks between sample groups, tabulation of institution-level characteristics for all sites, and a summary of attrition data for each site; Chapter IV recommends which test instruments should be retained or deleted, and discusses projected attrition to all sites. Appendices include written and tabular descriptions of the test measures, tester monitoring forms, flow charts of the analysis procedure, tables detailing response distributions, written definitions of behavior categories, tables of school characteristics, written and tabular attrition findings for each site, and subscales of the PDC Child Rating Scales. (CM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED160238

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

Project Developmental Continuity Evaluation

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

John M. Love

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

PS 010166

Pilot Year Impact Study: Instrument Characteristics and Attrition Trends

Interim Report IV, Volume 1

August 1976

This report was prepared for the Early Childhood Research and Evaluation Branch, Office of Child Development, Department of Health, Education and Welfare (Contract No. HEW-105-75-1114), Dr. Esther Kresh, Project Officer. Views or conclusions contained herein should not be interpreted as reflecting the official opinion of the sponsoring agency.

A PROCESS EVALUATION OF PROJECT DEVELOPMENTAL CONTINUITY

INTERIM REPORT IV, VOLUME 1:

PILOT YEAR IMPACT STUDY--INSTRUMENT CHARACTERISTICS AND ATTRITION TRENDS

August 1976

Prepared by:

Arthur C. Granville
Judy T. McNeil
Judy Meece
Sally Wacker
Mary Morris
Mel Shelly
John M. Love

With the Assistance of:

Barb Bruemmer
Jodi Bruemmer
JoAnn Emmendorfer
Ann Hale
Robert Hanvey
Dorothy Kelly
Helen Kiddon
Nancy Naylor
Jane Oden
Cathy Peterson
Leslie Ryan
Lynn Spencer
Jana von Fange

High/Scope Educational Research Foundation
600 North River Street
Ypsilanti, Michigan 48197

Table of Contents

Page

I. INTRODUCTION.................................... 1

An Overview of Project Developmental Continuity...... 1

Purposes of the PDC Evaluation..................... 3

Purposes of this Report............................ 3

Changes in Evaluation Plan Since the Last Report..... 8

II. METHODS.. 9

Data Collection Procedures......................... 9

Data Analysis Procedures........................... 15

Step 1: Does the Internal Consistency Coefficient Indicate Reliability?............................ 15

Step 2: Is the Internal Consistency Coefficient Constant from Fall to Spring?...................... 17

Step 3: Does the Test-Retest Coefficient Indicate Stability?.................................... 17

Step 4: Are the Measures Valid?.................... 18

Step 5: Are the Measures Sensitive to Change?...... 20

Step 6: Are the Measures Related to Social Competence?...................................... 22

Other Analyses..................................... 24

III. FINDINGS...................................... 27

Descriptive Characteristics of the Samples.......... 27

Reliability of the Instruments for Head Start Children.. 29

Validity of the Instruments for Head Start Children.. 33

Table of Contents
(continued)

Page

Sensitivity of Child Measures to Change.............. 37

Suitability of the Instruments for Use in the Higher Grades.................................... 43

Factor Structure of the Battery..................... 49

General Analysis of the PDC Classroom Observation System.. 58

Site Characteristics................................ 78

Comparability of PDC and Comparison Groups in Michigan.................................... 78

Recommended Modification of the Evaluation Design for Georgia................................ 78

Descriptive Characteristics of Centers and Schools.. 81

Analysis of Attrition Trends........................ 82

IV. CONCLUSIONS.................................... 89

Summary of Findings with Regard to Instruments....... 89

Justification of Recommendations................... 90

Additional Instrumentation Recommendations.......... 92

PDC Classroom Observation System.................. 92

PDC Child Interview (Faces)....................... 93

Preschool Productive Language Assessment Task (PPLAT).................................... 94

Summary of Findings with Regard to Attrition......... 95

APPENDIX A: Descriptions of the Measures in the Spring Battery.................................. 97

APPENDIX B: Forms for Weekly Tester Monitoring......... 107

Table of Contents
(continued)

Page

APPENDIX C: Flow Charts for the Analysis Procedure...... 123

APPENDIX D: Response Distributions for Child Measures... 131

APPENDIX E: PDC Observation Behavior Categories and Relative Frequencies of Behaviors Observed Using PDC Observation System............... 201

APPENDIX F: Center and School Characteristics.......... 225

APPENDIX G: Attrition Findings for Each Site........... 239

APPENDIX H: Composition of Subscales for the POCL and the CRS, Based on Factor Analyses.......... 257

I

INTRODUCTION

An Overview of Project Developmental Continuity (PDC)

The Office of Child Development originated Project Developmental Continuity (PDC) in 1974 as a Head Start demonstration program "aimed at promoting greater continuity of education and comprehensive child development services for children as they make the transition from preschool to school." The single most important effect of this undertaking, it is hoped, will be to enhance the social competence of the children served--that is, to increase their everyday effectiveness in dealing with their environment (at school, at home, in the community, and in society). PDC also aims to bring about broader and more intensive involvement of parents and teachers in the governance of school affairs, and to promote positive change in the institutional process, even beyond the people who may occupy the institution at a given time.

As part of the overall Head Start Improvement and Innovation effort, PDC emphasizes the involvement of administrators, classroom staff, and parents in formulating educational goals and developing a comprehensive curriculum. The object of this effort is to ensure that children receive continuous individualized attention as they progress from Head Start through the early primary grades. Existing discontinuities between Head Start and elementary school experiences will be reduced, if the program is successful, by PDC mechanisms that encourage communication and mutual decision-making among preschool and elementary school teachers, administrators, and parents.

Two program models provide alternative ways of establishing the administrative structure for continuity. In the Preschool-School Linkages approach, administratively separate Head Start and elementary programs are brought together by the device of a PDC Council, whose membership includes teachers, parents, and administrators from both organizations. In the Early Childhood Schools approach, Head Start and elementary programs are combined both administratively, by the Council, and physically, in the same building, creating a new institution. In both approaches a qualitatively different program is expected to emerge as a result of the Head Start-elementary school cooperation.

Continuity is expected to be established in two contexts: that of the individual child and that of the school structure. In the first context, continuity means, for example, that a child should not have to have his or her personal nature and needs rediscovered each year as he or she moves from one grade to the next; instead the child should become a more and more fully recognized member of the school "family" as time passes. In the context of school structure, continuity implies cooperative pursuit of common goals, and this involves articulation of philosophies and methods in all the various areas of school enterprise. It is expected that structural continuity will contribute directly to continuity in the attention given to individual children.

School organizations at 15 sites around the country received OCD funding during 1974-75 (Program Year I) to design and plan future implementation of the seven prescribed components of PDC. The components focus on:

- Administration: administrative coordination between and within Head Start and elementary school;
- Education: coordination of curriculum approaches and educational goals;
- Training: preservice and inservice teacher training and childrearing training for parents;
- Developmental/Support Services: comprehensive services (medical, nutritional, and social) to children and families;
- Parent Involvement: parent participation in policy-making, home-school activities, and classroom visits or volunteering;
- Services for the Handicapped: services for handicapped children and children with learning disabilities;
- Bilingual/Bicultural and Multicultural Education: programs for bilingual/bicultural or multicultural children.

During Year II, 1975-76, 14 sites (one had withdrawn voluntarily), comprising a total of 42 Head Start centers and elementary schools, began to implement PDC according to the plans they had drawn up during Year I, pilot-testing their adaptations of the program. At the end of Year II another site dropped out of the program.

In Year III, 1976-77, PDC is expected to exist in mature form at the 13 participating sites, and a decision will be made during that year, on the basis of the evidence presented, to maintain or modify OCD support for the entire demonstration program. The decision will be based in large part on consideration of the feasibility of evaluating PDC's effects on children's development over a long term. If the program is continued, it will be for a five-year period, from 1976 to 1981, during which its effects on children will be observed as the children in the focal group progress from Head Start through grade 3.

Purposes of the PDC Evaluation

The major purpose of the PDC evaluation is to aid the Office of Child Development in its efforts to design effective programs for early childhood education. To accomplish this, the evaluation will ultimately have to provide answers to the following critical questions about PDC's impact:

- How does PDC affect children's social competence?
- How does PDC affect parents?
- How does PDC affect the attitudes and workstyles of teachers and other staff?
- How does PDC affect the school organization in terms of philosophy, methods, and social climate?

In addition to describing the consequences of PDC, the evaluation will describe and analyze the processes that led to those consequences. Figure 1 illustrates the proportions of the total evaluation effort that are devoted to each component of the study. Although the assessment of child social competence is very important and is emphasized in the present report, the relationship of this to the rest of the evaluation should not be neglected. Volume 2 of Interim Report IV delineates the process evaluation more fully; it is sufficient to emphasize here that the aims of the total evaluation are to produce conclusions about what happened (impact) and how and why it happened (process). This information will facilitate future decisions about whether the program should be replicated, and if so, how replication can best be accomplished in the light of past experience.

Purposes of this Report

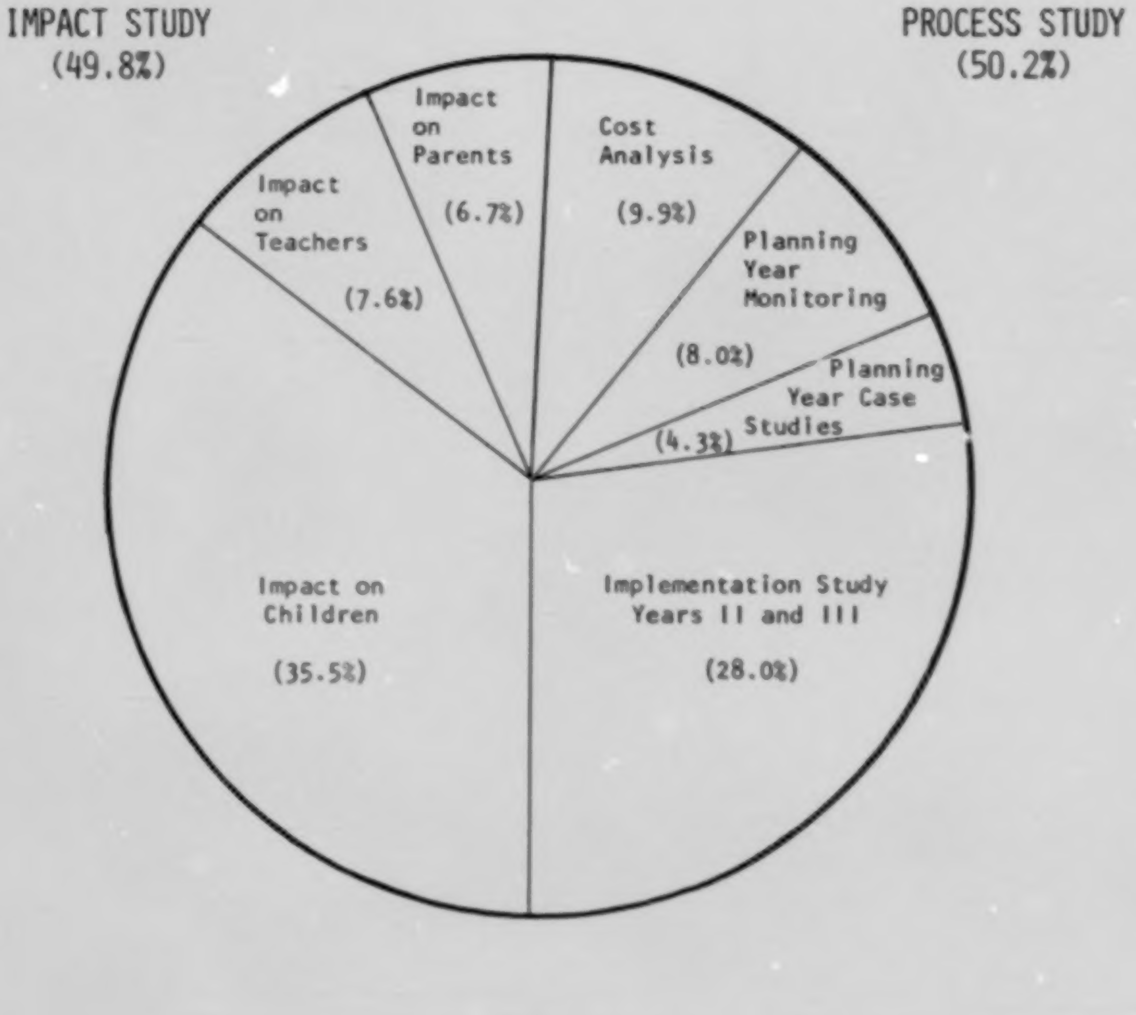
The present year, Program Year II, was reserved as a time for sites to try out and refine the program strategies they had developed during the planning year. There was no analysis at all of program impact this year--analysis of impact will begin with Program Year III in the coming fall.

10

Figure 1

PDC EVALUATION EFFORT

(Total 3-Year Study)



During 1975-1976 the evaluation methodology was also pilot-tested. Interim Report III, Part A (March 1976) addressed questions about three issues that are fundamental to the integrity of the evaluation:

(1) Are the measuring instruments appropriate to the task?

(2) Are the PDC and comparison groups really comparable?

(3) Will large enough samples of children remain in PDC and comparison schools at each site to permit a longitudinal study of program effects?

The present report extends the answers offered in the last report to these three questions. (However, there is greater emphasis this time on instrument suitability and less on group comparability.)

Instrument appropriateness. Since the ultimate goal of PDC is to enhance the social competence of children, it is vital to the evaluation that the instruments used actually do yield valid measures of social competence--valid not just in some theoretical sense, but in a way that can be demonstrated among the children who make up the population served by PDC.

The first step in establishing a test's validity is to establish its reliability, since an instrument cannot be valid unless it is first reliable; that is, it cannot yield relevant measures of a specified trait unless it yields consistent measures. For example, if each child in a classroom is given a certain language development test on Monday and then again on Wednesday, and the scores for the two testings are in wide disagreement, then one or both of the scores must not reflect language development truly because the reflections are inconsistent.

It is possible, though, that the test in question could yield consistent measures for other children (e.g., in a community where children are less inclined to be shy about interacting with the unfamiliar adult who administers it). Thus reliability and validity are not strictly inherent properties of an instrument, but may vary from sample to sample. Accordingly, in our last report reliability and validity data were examined and tabulated separately for each site. However, although differences were found in reliability and validity coefficients from site to site, there was no indication that these were other than random differences. Thus in the present report the analysis of test characteristics is based upon data pooled across all sites, on the grounded assumption that the tests "behave" in about the same way at every site. The aggregate sample, in other words, is believed to be representative of the samples at each separate site.

There are qualities beyond present indications of reliability and validity that affect an instrument's suitability for inclusion in the battery. The other qualities examined during this analysis period included each instrument's relation to a hypothetical "social competence" criterion, its relation to all the other instruments in the battery, its sensitivity to change in the trait measured, its appropriateness for children up to grade 3, and the ease with which it can be administered. The findings produced by analyses of these instrument characteristics are all included in this report in response to question 1 above.

Group comparability. Last fall, the number of children in the Michigan site for whom testing was completed was too small to permit the analytic contrast of PDC vs. comparison group characteristics that was performed for samples at other sites. Therefore this analysis was carried out on data collected in Michigan this past spring, and its results are contained in the present report. Since the fall comparability analyses were based only on measures that were found to be reliable and valid at each respective site, the same procedure was observed in the spring analysis for Michigan--only the measures that met minimum reliability and validity criteria there entered into the comparability check.

Because the last reporting deadline did not permit presentation of institution-level data that recently had been gathered for all PDC and comparison Head Start centers and schools, these data, collected in fall 1975, are included now in this report. Since PDC and comparison schools were originally selected on the basis of similarity, and since continuing efforts are being made at each site to match the two groups of children on background variables, it was not judged necessary to evaluate the comparability of institution-level data. Therefore the figures are offered simply for purposes of description and for possible later use as explanatory variables.

Projected attrition. In early spring, each site coordinator was asked to identify children who in past years had been enrolled in what are now PDC and comparison centers, and to determine how many of the same children are currently enrolled in PDC and comparison elementary schools, respectively. The purpose of gathering these figures was to obtain indications of the rate of attrition among children at each site who were the predecessors of the children in the present Head Start classes. Projections were made, based on attrition rate calculations, of the number of children in present Head Start samples likely to still be in PDC or comparison schools, as appropriate, come third grade (the point at which the prospective longitudinal study will end). The results of this attrition study are given here in tentative answer to the question of whether a sufficient number of children will remain to permit a conclusive longitudinal study.

Included also in this section is a discussion of the strategy proposed for the Georgia site, which, instead of requiring a contemporaneous comparison group, involves comparison with a cross-section of children who are in grades K, 1, 2 and 3 during the PDC group's Head Start year. In each later year, the PDC children's scores are compared with scores predicted on the basis of trends observed among the upper-grade children.

The contents of the chapters to follow are described briefly below.

Chapter II, Methods. The data collection and data analysis procedures that were followed in the spring are summarized here.

Chapter III, Findings. This chapter presents:

- Findings on the reliability, validity, and general utility of the instruments included in the spring battery;
- Results of the check on comparability of the PDC and comparison groups in Michigan;
- A tabulation of institution-level characteristics for all sites;
- A summary of attrition data for each site, along with projections of the number of children from the original groups likely to be remaining at the end of grade 3.

Chapter IV, Conclusions and Recommendations. We present here conclusions and implications drawn from our findings on the major issues: adequacy of the measuring instruments, comparability of groups in Michigan, and projected attrition at all the sites.

Changes in Evaluation Plan Since the Last Report

Changes in the sample of sites. Since the last report, OCD decided that a modified evaluation strategy would be more appropriate for the Arizona site because few tests could be located that were suitable for use among Navajo-speaking children as measures of child impact. In addition, the enrollment projected for the PDC and comparison groups in 1976-77 was much smaller than the number recommended, so that even if development of original Navajo-language tests had been undertaken, it is unlikely that the sample available would have been adequate for evaluation purposes. The Navajo site will instead be evaluated through a case history approach that will document program process and perceived impacts.

New Jersey's participation in PDC has been discontinued entirely; the program there is no longer receiving OCD funds.

Postponement of surveys. Administration of the Teacher and Parent Surveys at each site had been scheduled for the spring of this year, and the results of the surveys were to have been presented in this report. The surveys were postponed, however, because the required clearance has not yet been received from the Office of Management and Budget.

Postponement of PPLAT analysis. An analysis of the Preschool Productive Language Assessment Tasks (PPLAT), an instrument that was administered on a trial basis at two sites this spring, had originally been planned for this report. However, the scoring of this test is complex, requiring development of a special computer program, and the program was not completed in time to permit analysis of PPLAT data. Presentation of the results will be deferred until the next report.

II

METHODS[1]

Data Collection Procedures

To maximize the potential for collecting high-quality data the following procedures were initiated:

1. An organizational structure for individuals involved in the data collection effort was outlined and role responsibilities were defined.
2. A training model was designed that specified tester performance standards and provided for large-group, small-group and individualized instruction, daily analysis of individual tester performance, and discussion of potential problems.
3. On-site monitoring of testers by trainers was conducted prior to the start of the actual testing.
4. Weekly monitoring was done by site interviewers.
5. A weekly check of all currently completed data was made by site coordinators.

Each of these procedures is discussed below.

Field organization. In order for the data collection to be systematic and organized, role responsibilities were explicitly outlined. For example, site coordinator responsibilities included contacting the PDC coordinator regarding the start of testing, setting up and chairing a meeting with the Head Start teachers involved in the evaluation, keeping in contact with the supervisor of field operations about the status of data collection and any problems that the site was having, checking all completed data on a weekly basis, keeping up-to-date records on the status of the data collection, carrying out any needed training, observing and testing children, and monitoring testers.

---

[1] The tests and other instruments used, and the order of their administration, are described in Appendix A. Further details on testing, monitoring and other procedures followed by testers on site can be found in the Field Procedures Manual, March, 1976.

Of the 45 testers involved in the fall testing, 26 returned in the spring and 12 new testers were hired. For the most part, the additional hiring was done in those four sites (Georgia, California, Texas and Michigan) where both PDC and non-PDC Head Start children (or, in the case of Georgia, the elementary school children) were included in the spring testing sample. If there were at least two returning testers in a site (other than those mentioned above) no new testers were hired since it was feasible for two people to test the PDC Head Start children in 6 weeks. Thus, the number of testers per site ranged from 2 to 5.

Training model. Two intensive training sessions were held; a four-day session for the nine tester-trainers in early March and an eight-day session for the 38 testers in mid-March. The methods of training were similar for both groups and involved explanation and demonstration of a measure to the entire group, followed by practice in small groups. During the training session for testers, the initial practice session after the large group introduction and demonstration involved the use of test "scripts." The scripts consisted of test instructions, child responses and rationales for scoring. In using the scripts, two testers would pair up and one, the "child," would perform as indicated on the script while the other tester administered the test without the script. This provided an excellent learning situation since the child responses included on the script covered all the administration rules and gave the testers a chance to work with and correct each other without having to have a trainer nearby to answer all their questions. Two scripts were written for most of the tests.

Rather than have the new testers learn both the observation system and the child tests (which is too much for inexperienced testers to cover completely, even in 6 full days) the decision was made to have them focus solely on the tests and master their administration. Thus, there were essentially two different training sessions going on simultaneously--one for the experienced testers which included the observation system, and one for the new testers which provided more practice time. When possible, the two sessions were held jointly (as for introduction and demonstration of a child measure) to give the experienced testers and the new testers the opportunity to work together and learn from one another.

The revised classroom observation instrument was introduced to the returning testers as a group with trainers demonstrating the behaviors reflected on the instrument and returning testers identifying the observed behaviors. After the PDC observation manual was reviewed and discussed in small groups, the testers, in groups of two and three, wrote scripts incorporating behaviors on the observation system. They then acted out the scripts while the other testers coded the behaviors. A better understanding of the behavioral categories and items was gained by this script-writing and acting. Videotape reliability data were collected toward the end of the training session (the videotape had audible time signals on it that corresponded to the beep tapes used by testers in the classroom). However, the trainees had a difficult time seeing and hearing the target child and coding the behaviors. Videotapes were used minimally during the training process because of these problems.

Since it is critical that interviewers administer the tests in a standard way, each interviewer was systematically "checked-out" on all of the child measures during the training session. During this procedure a trainer played the role of the child (also recording the "child's" responses) while a tester administered one or more of the child measures to her. Prior to these "check-outs" the trainers had decided how the trainer (acting as the child) would respond to each item on each test. This was done for two reasons: 1) to insure that each interviewer was exposed to the same situation and 2) to incorporate child responses that covered all test administration directions. For example, on the PIPS interview, there are specific things for the tester to say if the child gives an unrelated answer, a repeated answer, refuses to answer, etc. By exhibiting all these behaviors in the check-out situation, trainers were able to assess the tester's understanding and expertise in administering each of the child measures.

Standards were set for acceptable performance during these check-outs and if they were not met, additional training and practice were prescribed. Check-outs were then repeated at a later time during the training session to insure correct test administration.

The trainers met every evening to discuss the day's activities and to report on the progress or status of each tester. Potential problems were identified and discussed as were necessary schedule changes.

On-site monitoring. In all sites except one the data collection started one week after the training session (Iowa had a spring break and so the testing started two weeks after training in that site). During this interval, testers practiced administering the tests together, and where possible, administered the tests to preschool-aged children not in the sample. They also met with the PDC Coordinator, the Head Start teachers involved, and spent time in the classrooms acquainting or reacquainting themselves with the children.

During the weeks of March 24 and April 5 a trainer visited each of the sites for two days to monitor each of the interviewers. The purpose of this visit was threefold: 1) to monitor each tester as she/he administered the measures to a child, 2) to review thoroughly the observation system and observe in the classroom along with the testers (for reliability purposes) and 3) to review the monitoring system with them as a group so that they would be able to monitor each other on a weekly basis.

Although it was not feasible for the testers to administer the tests to children while they were in Michigan, it was important that each tester have such experience before starting the data collection. Thus, during the on-site visit the trainer watched each tester administer the sessions and provided feedback on his/her interactions with the children. This experience served two purposes; it gave the trainer an indication of how individual testers would establish rapport and interact with children and it helped alleviate some of the anxieties the inexperienced testers felt about administering the measures to children.

Weekly monitoring. During the course of a testing week, testers alternately monitored each other; the one acting as monitor simultaneously completed the test booklets and the individual monitoring forms for each test. After the session, the monitor and tester discussed any errors and the monitoring booklets and forms were sent to the High/Scope Foundation. The monitoring forms are shown in Appendix B. The categories on the forms beside which an X appears are those in which testers, as a group, made more errors than expected or than was judged tolerable. Of the 110 categories across all 13 tests, there were only 5 with an error rate this high. These areas will be stressed during the next training session.

Weekly pre-transmittal data checks. Testers were required to give or send their completed data to the site coordinator at the end of each week. The site coordinator then checked these tests, plus any she/he had completed, for recording and scoring errors. (Site coordinators and interviewers had received a checklist specifying what to look for when reviewing each completed booklet, e.g., "Is the identification complete?", "Did the interviewer fail to give a second trial when it should have been given?", "Did the interviewer skip an item?", etc.)

Errors were pointed out to the particular tester and, if necessary, further training was provided by the site coordinator. The site coordinator also kept track of all completed data (in addition to the individual records each tester kept of his/her classes) and mailed the completed data to the High/Scope Foundation on a weekly basis.

Recording and scoring of data. Data received from each site at the High/Scope Foundation were checked by the Supervisor of Field Operations. Errors in recording or coding were identified and explained to the site coordinators, who then discussed them with the other testers.

Once the raw data had been screened for accuracy at High/Scope, the test protocols were tagged with unique identification numbers for each student, scored and verified, then keypunched and verified.

Data collection sequence. In those three sites where the data were collected on both PDC and comparison children, the same procedure was followed as in the fall: each tester was assigned to test children in both the PDC and comparison classes and the data on children in both groups were collected simultaneously. At the remainder of the sites only the PDC children were tested, and testers, for the most part, were assigned to classes they were familiar with.

The children were listed on the testing rosters in the order of their fall testing date and testers were instructed to test children in that order (e.g., the children tested first in the fall would be tested first in the spring). This was to maintain a test-retest interval that would be fairly constant for all children. The mean interval between fall and spring testing was 6.8 months, with a standard deviation of .8 months.

The procedure followed by testers in determining the language capabilities of children in the testing sample was (1) to ask each child's classroom teacher for his or her judgment, (2) to observe the child's verbal behavior under natural classroom circumstances, and (3) on the basis of these indications, to administer the English or Spanish versions, or both, of the Bilingual Syntax Measure (BSM). In most cases, all three steps of this procedure produced consistent conclusions, and subsequent testing was accordingly conducted in English or Spanish or both. (In some cases, this screening process led to the conclusion that the child was proficient in some third language, but not English or Spanish. These children were deleted from the testing sample.) When the conclusions produced were inconsistent, the tester made a judgment of the child's language capabilities based upon all the available information.

The spring PDC data collection started the week of March 28th in all the sites except Iowa, where it started a week later. The length of the data collection period was fairly constant across the sites, with most finishing up within six weeks. The data collection effort progressed much more smoothly during the spring due to tester familiarity with the tests, the testing procedure, the children, the teachers, etc., and was successfully completed in all sites.

Data Analysis Procedures

The primary purpose of the spring analysis of child measures[1] was to determine their adequacy for use in the longitudinal evaluation of PDC. Secondarily, data from the Michigan site were analyzed to determine the comparability of PDC and comparison groups there, since sufficient data had not been produced during fall testing to permit this comparison. For both purposes it was necessary to examine (or re-examine) the reliability and validity of the instruments included in the spring battery. The primary analysis, which focused on psychometric characteristics of the instruments, consisted of six major steps, as shown in Figure 2. These steps dealt sequentially with:

- internal consistency in spring 1976;
- constancy of internal consistency from fall to spring;
- test-retest (fall-spring) reliability;
- validity;
- sensitivity to fall-spring change; and
- relevance to social competence.

The details of these steps are described in the following text. Flow charts depicting each of the six steps separately can be found in Appendix C.

Step 1: Does the Internal Consistency Coefficient indicate Reliability?

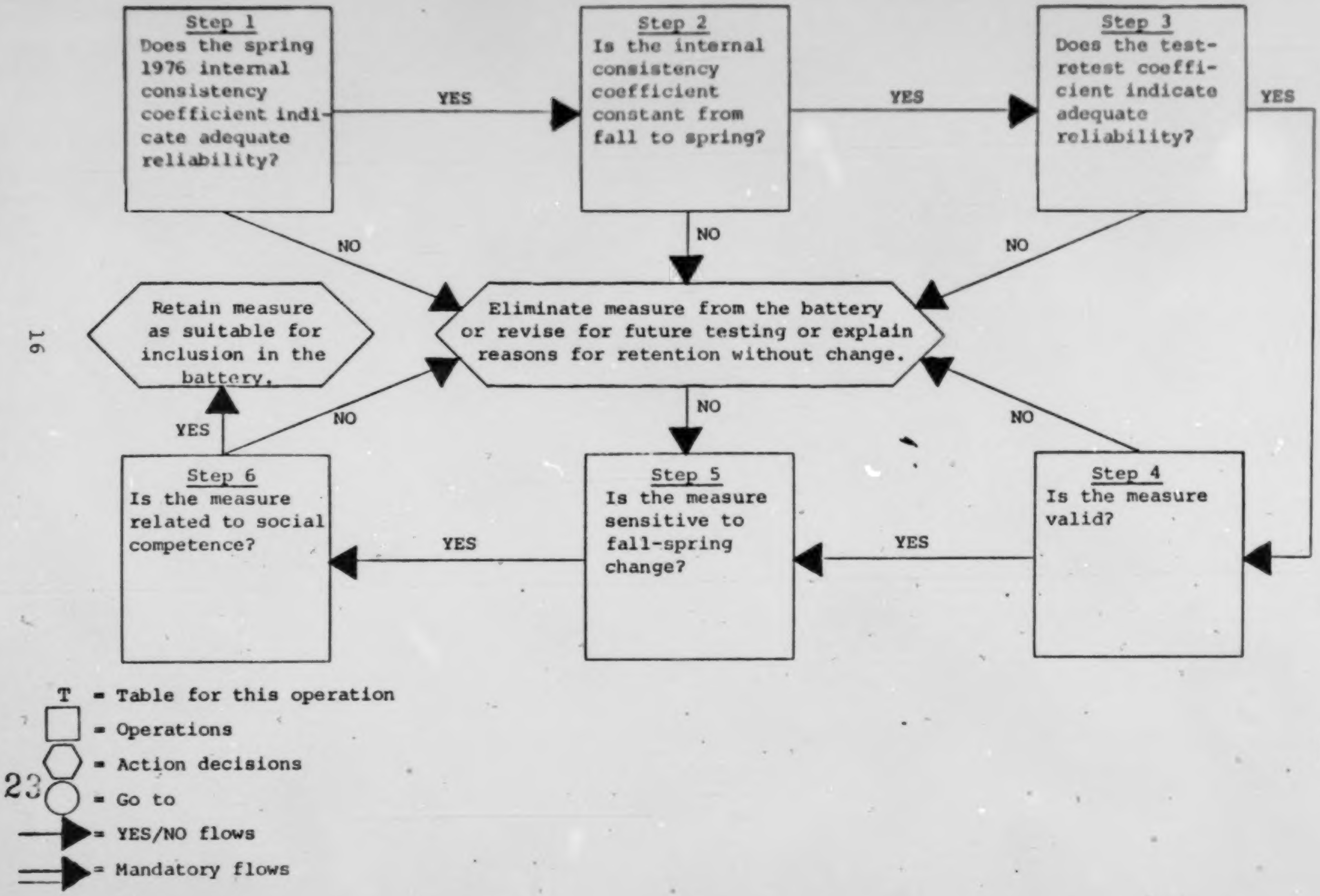
Cronbach's alpha, a measure of internal consistency, is an index of the amount of overlapping variance among the items that comprise a scale; it may be conceived as the mean of all possible split-half correlations (e.g., a correlation of the odd and even items within a scale). Coefficient alpha can be interpreted as a measure of the reliability of an instrument at a single administration since it reflects the degree to which the constituent items of a test relate to a unitary construct--the degree to which they tend to measure the same thing.

---

[1] The procedures described in this section refer to all child measures except the Preschool Interpersonal Problem Solving Test and the Classroom Observation System. Internal consistency analysis is inappropriate for the PIPS so Steps 1 and 2 were omitted from the analysis sequence for that measure. The analysis of the Classroom Observation System is treated separately in Chapter III.

Figure 2

Master Flow Chart for Analysis of Instruments in the Spring 1976 Battery


Step 1
Does the spring 1976 internal consistency coefficient indicate adequate reliability?
YES
Step 2
Is the internal consistency coefficient constant from fall to spring?
YES
Step 3
Does the test-retest coefficient indicate adequate reliability?
YES
NO
NO
NO
Retain measure as suitable for inclusion in the battery.
Eliminate measure from the battery or revise for future testing or explain reasons for retention without change.
YES
NO
NO
NO
Step 6
Is the measure related to social competence?
YES
Step 5
Is the measure sensitive to fall-spring change?
YES
Step 4
Is the measure valid?
T = Table for this operation
= Operations
= Action decisions
= Go to
= YES/NO flows
= Mandatory flows

The procedure followed for the determination of spring 1976 internal consistency reliability is pictured in Figure C-1, Appendix C: internal consistency was determined for each measure for the two combined samples of English-dominant and Spanish-dominant children taking each test.[1] If the alpha for a measure was over .65 within either language sample, the measure was considered internally consistent for that group for the spring test administration; if the alpha value for a measure fell below .65, that measure was likely to be dropped from the analysis procedure.

Step 2: Is the Internal Consistency Coefficient Constant from Fall to Spring?

Figure C-2, Appendix C, outlines the procedure for comparing fall and spring Coefficient alpha values. For each measure, the spring alphas for the English and Spanish language samples were compared to the fall alphas for the English and Spanish "best aggregates."[2] Measures that were reliable at both time points (alpha $\geq$ .65) were considered to have demonstrated adequate internal consistency and are likely to be retained in the final test battery. Problems with internal consistency at each time point are discussed in the Findings section for each particular measure. Measures that failed to achieve an alpha value of .65 at both time points are likely to be dropped from the battery.

Step 3: Does the Test-Retest Coefficient Indicate Stability?

Having data available from two administrations of these child measures for the same children, it was possible to assess the reliability over time (stability) of the measures for individuals and for the two language groups. Figure C-3, Appendix C, illustrates the procedure followed for fall-to-spring reliability checks. Stability of the measures for individuals was assessed using the Pearson product-moment correlation as the coefficient of test-retest reliability. Since, in most cases, the time between test administrations was relatively short, it seemed logical that spring scores would correlate strongly and positively with fall scores for individual students. On the other hand, the interval (about six months) was greater than usually desirable for test-retest reliability--especially considering that it represented six months during a child's first preschool experience. Therefore, extremely high correlations were not expected nor were they required to demonstrate reliability. The criterion for this measure of individual stability was a $\underline{r}$ value greater than .40.

---

[1] The English-language sample includes all English-dominant, non-handicapped students for whom each measure was obtained. The Spanish-language sample includes all non-handicapped, Spanish-dominant students for whom each measure was obtained in Texas and California. (The procedure for deciding which language would be most appropriate for testing bilingual children was described on p. 13.)

[2] In addition to meeting the criteria outlined for English language and Spanish language samples, aggregate samples for the fall alpha computations were selected so that only those students who had non-missing data on all reliable and valid tests were included. The English aggregate included only those sites where test scores of PDC and comparison children were comparable.

Step 4: Are the Measures Valid?

The procedures followed for the determination of validity of measures are pictured in Figure C-4, Appendix C. As with reliability, previous research provides some information on the validity of the measures. But the validity of the measures also must be ascertained within the context of the PDC evaluation. Most of the measures were selected from larger existing batteries, and items on most of the measures have been modified, both to meet the needs of the sample being tested and to permit use by paraprofessional testers. Therefore the validity of the measures within the PDC environment, and within the test battery in which they are administered, must be determined anew. The concern in this report is with concurrent validity--a measure's relationship to other measures of the same construct and to measures of other constructs: a measure should correlate highly with other measures of the same constructs, and should not correlate at all with measures of independent constructs.

An hypothesized correlation matrix was constructed prior to the fall data analysis, based on knowledge of the constructs the measures were presumed to represent. The values in the matrix indicate the level of relationship that theoretically should obtain between the measures if they are valid measures of the constructs. The actual spring correlations (within language groups) were then evaluated against the hypothesized correlations.

The hypothesized correlation matrix was constructed by determining first the correlations within the three areas of child tests; that is, within Cognitive-Language measures, within Psychomotor measures, and within Social-Emotional measures. Then the desired correlations among the three groups of tests were determined. Generally, higher correlations were expected within an area than between areas. But since each area is actually composed of linked constructs rather than alternative measures of the same construct, very few high correlations were expected.

The actual spring correlations between measures (the ones found reliable) were calculated within each language group, and the following procedure used to determine whether a given measure was valid. First, the obtained intercorrelation matrix was compared with the hypothesized matrix and deviations of each correlation from the hypothesized one were calculated (e.g., if the hypothesized correlation was "medium" and that obtained was "low," a deviation of "-1" was scored; if the hypothesized correlation was "zero" and that obtained was "medium," a deviation of "+2" was scored). For each measure, the absolute values of the deviations were summed across all measures and divided by the number of measures. If this ratio had a value of 1.0 or less, the measure was considered valid. The criterion implicit in this procedure is that a measure's concurrent validity is adequate if, on the average, the obtained correlations with other measures are within the range adjacent to the expected value.

This procedure allows for rather large deviations from the hypothesized relations, but the procedure was judged to permit a useful first approximation to establishment of validity.

For each measure, the actual spring correlation matrix of reliable measures for the English- and Spanish-language samples was compared with the actual fall correlation matrix of these same measures for the English and Spanish "best aggregates." In determining the validity of a measure, most weight was given to the comparisons of actual spring values with the hypothesized conceptual framework. However, if a measure failed at both time points to fall within the hypothesized range, but fell within the same range (e.g., "medium") at both time points, consideration was given to the possibility that these obtained values might more closely represent the actual degree of relationship between measures to be expected in the PDC populations (see the Results section for a discussion of specific measures.)

Step 5: Are the Measures Sensitive to Change?

An important criterion of an instrument's utility is whether or not the scores it produces reflect change when change can be assumed to have occurred. Figure C-5, Appendix C illustrates the procedures followed to assess sensitivity to change for each measure for the PDC English- and Spanish-dominant samples. For all non-handicapped students with non-missing data at both time points, paired t-tests were computed comparing fall and spring scores on each reliable and valid measure. Spring means for groups were predicted to be greater than fall means. Each measure for which the spring mean was significantly greater than the fall mean (using a directional t-test) was considered sensitive to gross changes in groups over time. However, since with sample sizes as large as these slight changes in mean scores might produce a significant result, further analyses of sensitivity to change were necessary to assess the suitability of the instruments as change measures in a longitudinal analysis. These further analyses were based upon consideration of the separate sources of change in the behavior of children in the PDC testing samples: principally these sources are maturation and the school program. Program-related change was not expected to be detectable from fall 1975 to spring 1976 because of the shortness of the interval and the relatively small differential between PDC and comparison programs at the Head Start level. Thus the analyses focused on age-related change, as will this discussion.

There are two general ways of viewing the age-relatedness of a test score: At a single time point, for a group of children, differences in the children's ages may be related to differences in their test scores. The correlation between age and test score is a measure of the strength of this relationship. In this sense, age is a characteristic each child has in relation to other children. A researcher can predict the difference between the test scores of two children based on some proportion of the differences in their ages. For a particular child, the difference between his or her test scores at two time points may also be related to the difference in age at those two time points. In this sense, age is a characteristic each child has in relation to himself or herself. With other status variables controlled, the relationship of an individual child's age to test scores over time should be similar to the relationship between age and test score in a group of mixed-age children at a single time point.

Between the fall and spring testings, the child has changed in age. It is logical to expect that if the PDC child measures are developmental (i.e., sensitive to age-related changes), child A, whose age was 50 months in the fall and 55 months in the spring, should have a spring score that is similar to child B's fall score if child B was 55 months in the fall. It is also reasonable to expect that five months of educational experience should have added something to a child's score over and above mere aging. So, it made sense to predict that child A in the spring would do as well as (and probably better than) child B did in the fall if they were matched on other status variables (or if these were statistically controlled).

The following strategies were employed in assessing age-related change:

1. Fall child measures were regressed on child status variables, including fall age. The overall R^2 and the significance levels of each predictor were examined.

2. Spring child measures were regressed on child status variables and spring age. The overall R^2 and the significance levels of each spring predictor were compared to the fall predictors. This replication was necessary to determine whether or not these predictors stood in the same relationship to the child measure as they did in the fall. Unmeasured factors, such as the difference between programs or levels of implementation, or increasing differences between ethnic groups or previously treated (preschool) children, might have seriously altered the predictive power of the status variables measured. If fall and spring regression equations were not comparable, the analysis was terminated here.

3. Based on knowledge of status variables for each child in the fall sample, expected spring scores were computed using a multiple linear regression prediction model, substituting the new age (spring) for the old age (fall) in the fall prediction equation (separately for English and Spanish language samples).

Example:

Fall Score: BSM-E = a_0 Unit Vector + a_1 Fall Age + a_2 Sex + a_3 Ethnicity + a_4 Program + a_5 Preschool + a_6 Siblings

Expected Spring Score: BSM-E = a_0 Unit + a_1 Spring Age + a_2 Sex + a_3 Ethnicity + a_4 Program + a_5 Preschool + a_6 Sibilings

Within language groups, spring *observed* mean scores were compared with spring *expected* mean scores. (In addition, a status variable was created that classified children on whether or not their *observed* score was higher than, equal to, or lower than the *expected* score. The percentage of children having scores higher than or equal to fall scores was predicted to be larger than the percentage whose spring scores were lower than their fall scores.)

If the mean *observed* spring score was greater than or equal to the mean *expected* spring score, the measure was considered sensitive to age-related change. If the observed score was less than the expected score, the measure was not considered sensitive to age-related change for this testing interval. Nevertheless, the measure might be retained for future testing since several other factors might be influencing performance:

- The measure might have been designed to tap relatively stable child dimensions, and as such may be used as a covariable in later analyses.
- The child of 4 or 5 may not gain competence on this task during this age period, but may show significant maturity and educational experience-related gains at later ages.
- This measure might be sensitive to changes over greater time periods (e.g., 12 months or 18 months).

Step 6: Are the Measures Related to Social Competence?

If social competence is not simply an unconnected collection of specific abilities but a fairly global trait corresponding to something like "know-how" or "savvy," then we might reasonably expect each individual measure in the battery to be empirically related to a central criterion of social competence. To test this assumption, the ratings assigned to each child in the sample by his or her teacher (using the PDC Child Rating Scale) and those assigned by the tester (using the Pupil Observation Checklist) were selected as the most general available measures of "social competence[1]," and the relationship of the other measures to these was examined. The rationale for the selection of these measures was that a broad range of the child's characteristics--social, emotional, cognitive, linguistic, and psychomotor--is spanned when the teacher and tester assess the degree to which the child...

---

[1] When speaking of a *measure* of "social competence," as distinct from the *concept* of social competence, humility dictates the use of quotation marks, since we are far from being able to assume that what has been measured is a genuine embodiment of social competence.

"Has the desire to master all kinds of skills,"

"Expects to succeed, not afraid of failure,"

"Attempts to solve social problems with little adult assistance,"

is "Active vs. Passive,"

is "Quick to respond vs. Needs urging,"

"Keeps trying vs. Gives up easily."

There are 39 items in the rating scale completed by the teacher and 12 in the scale completed by the tester, assuring a fairly broad report of the impression the two have gained of the child's general performance.

Having selected these two sets of measures as "social competence" criteria, the strategy for examining the relevance to social competence of the other measures in the battery was to determine how well scores on each of these other measures could be predicted from knowledge of a child's status on the "social competence" criteria. The more accurately a test score could be predicted, the firmer the grounds for assuming that the trait measured by the child test is truly relevant to the development of social competence. It is important to note, though, that the reverse reasoning does not hold: the finding of a weak relationship does not necessarily mean that the trait measured by the child test is unrelated to social competence--it may be, for example, that the relationship is one that cannot be detected at an early stage of development, and it may also be that the "social competence" criteria selected do not adequately represent the concept of social competence.

The procedure followed in this exploration of each measure's relevance to social competence was first to factor analyze all the 51 items contained in the PDC Child Rating Scale and the Pupil Observation Checklist (the instruments completed by teacher and tester, respectively) to determine how the items might best be aggregated. On the basis of the factors that emerged, factor scores were computed for each child, then these scores and all other measures were factor analyzed together. (If the aggregated teacher and tester ratings truly represent assessments of competence in general rather than particularized competence, the factor scores created from the ratings would be expected to load prominently on the general factors found within the total set of child measures; thus this analysis provided a tentative check on the validity of the criteria.) Finally, each measure produced by the child tests was regressed on the "social competence" criteria. In other words, the criteria served as the predictor variables and each child measure was the variable whose value was to be predicted. The results permitted an examination of the relationship between each child measure and the "social competence" criteria, collectively and singly. Figure C-6, Appendix C, illustrates the analytic process just described.

Other Analyses

In addition to the analyses of test characteristics, two other sets of analyses were performed: examination of the comparability of the PDC and comparison samples in Michigan and examination of the suitability of the instruments for children in higher grades.

PDC-comparison group comparability in Michigan. In the fall, difficulties encountered in data collection made it impossible to gather the data needed to determine the comparability of groups in Michigan as had been done for all the other PDC sites. Thus both groups at this site were re-tested in the spring to permit this essential analysis. The procedure involved comparing the groups on the same demographic variables involved in the fall comparisons for other sites, and on the performance measures found to be reliable and valid for the Michigan sample this spring. For every variable, all available data entered into a test of the equality of PDC vs. comparison group status. For categorical data (on ethnicity, for example), the equality of PDC and comparison group proportions was evaluated by means of the chi square statistic; for metric data (all test scores), equality of group means was determined by t tests. The criterion of significance for each statistical test was a probability value of less than .10.

Suitability of the instruments for children in higher grades. In Maryland, 30 third-grade children were selected for testing this spring in order to permit examination of the psychometric characteristics of the battery for children at this upper extreme. (If the PDC evaluation is extended into a longitudinal study, the target cohort of Head Start children will be followed through their third-grade year.) In Georgia, where the cross-sectional comparison group consists of children in grades K - 3, spring testing of the comparison children served the dual purpose of providing the cross-section of scores required by this special design and also of yielding further indications of the tests' suitability for children in the upper grades.

The analysis procedure was essentially the same as that followed for the children in the Head Start samples except that no comparison of fall-spring data was involved; the objectives were to examine the tests' reliability, validity, score distributions, and relevance to social competence for these "harbinger" children.

III

FINDINGS

Descriptive Characteristics of the Samples

In fall 1975, PDC and comparison group children were tested at 14 sites. In the spring, testing was conducted at 13 sites (the decision had been made, as noted in Chapter I, to examine program impact in Arizona by alternative means), and at 9 of these sites, only the PDC groups were tested because, according to the original plan, children in the pilot year comparison groups were not to be followed longitudinally. However, in Michigan, comparison children were tested because it had not been possible to complete testing there in the fall; in Georgia, the cross-sectional (elementary-level) comparison group was tested in order to obtain baseline measures of children's spring status against which to compare the performance of the PDC children in later years; and in California and Texas, the two bilingual sites, comparison children were tested in order to ensure a Spanish-dominant sample of the largest possible size--note in Table III-1 that California and Texas are each divided into English-dominant and Spanish-dominant samples.

Children were eliminated from the analytic sample (the sample from which data were drawn for the various analyses) if they were identified as having a handicap likely to interfere with development or with testing (e.g., hearing impairment). Children with non-debilitating handicaps were retained, however. Children with debilitating handicaps will be included in some aspects of the evaluation in the future but are excluded at this stage from analyses that focus on test characteristics.

The first column in Table 1 shows the number of all children tested at each site; the last column shows the number of those retained in the analytic sample. For purposes of general information, the table also shows ethnic proportions and proportions of Spanish- and English-dominant children at each site. In the bottom row, these variables are summarized across all sites and all groups.

Table 1
Descriptive Characteristics of the Samples for Spring Data, 1976

		Number in Full Sample	Percent Handicapped	ETHNICITY						DOMINANT LANGUAGE			Number in Final Analytic Sample
				% Black	% Hispanic	% Amer. Indian	% White	% Asian	% Other	% English	% Spanish	% Other	
CALIFORNIA-Eng	PDC	31	0	10	80	03	03	03	0	100	0	0	31
	Comp	27	28	0	79	0	21	0	0	100	0	0	21
CALIFORNIA-Span	PDC	5	0	0	100	0	0	0	0	0	100	0	5
	Comp	11	9	0	100	0	0	0	0	0	100	0	10
COLORADO	PDC	37	42	0	82	0	18	0	0	100	0	0	25
CONNECTICUT	PDC	41	15	76	17	0	07	0	0	100	0	0	30
GEORGIA	PDC	34	28[a]	68	0	0	32	0	0	100	0	0	28
	Comp	111	22[a]	41	0	0	59	0	0	100	0	0	104
FLORIDA	PDC	21	0	100	0	0	0	0	0	100	0	0	21
IOWA	PDC	34	07	74	0	03	24	0	0	100	0	0	32
MARYLAND	PDC	45	05	52	0	0	41	02	05	100	0	0	42
	3rd Graders	30	0	53	0	0	47	0	0	100	0	0	30
MICHIGAN	PDC	32	03	66	03	0	31	0	0	100	0	0	32
	Comp	51	02	96	0	0	04	0	0	100	0	0	50
NEW JERSEY	PDC	45	05	98	02	0	0	0	0	100	0	0	43
TEXAS-Eng	PDC	12	0	0	67	0	33	0	0	100	0	0	12
	Comp	4	0	0	25	25	50	0	0	100	0	0	4
TEXAS-Span	PDC	28	0	0	93	0	07	0	0	0	100	0	28
	Comp	34	0	0	97	03	0	0	0	0	100	0	34
UTAH	PDC	31	17	04	32	0	64	0	0	100	0	0	27
WASHINGTON	PDC	41	23	37	08	05	50	0	0	100	0	0	32
WEST VIRGINIA	PDC	35	29	14	0	0	83	03	0	100	0	0	25
ALL GROUPS COMBINED		740	12	44	24	01	30	01	01	89	11	0	666

---

[a] These are revised figures, based on a review of handicapping conditions conducted after completion of the analyses presented in this report.

Reliability of the Instruments for Head Start Children

Two estimates of reliability were calculated for the child measures: internal consistency, as defined by Cronbach's alpha; and stability, the correlation between fall and spring scores. Table 2 shows the internal consistency of the measures, based on fall 1975 scores and on spring 1976 scores. Table 3 shows the stability of the measures from fall 1975 to spring 1976. A measure had to achieve an internal consistency alpha of at least .65 to be considered reliable for use with this population. A fall-spring correlation of .40 was considered desirable, but less stress was placed on this measure of stability because it spans a fairly long interval of time for test-retest reliability, especially for a child's first year of preschool experience. In addition, slight changes in test administration and/or test scoring were made in many of the tests between fall and spring and these changes could have reduced the correlations.

To summarize the reliability findings, the following tests are judged to be reliable for use in the PDC evaluation: BSM-E, BSM-S, WPPSI Block Design, Verbal Fluency, Verbal Memory, Arm Coordination, Draw-A-Child, POCL, PIPS, the Child Rating Scale, and the site-specific measures (Opposite Analogies and Do You Know...?). Each instrument is discussed in detail below.

Bilingual Syntax Measure-English. The BSM-E is considered to be reliable for English-dominant children. It achieved acceptable levels of internal consistency in the fall and spring and an acceptable fall-to-spring correlation for this sample. An insufficient number of Spanish-dominant children received this test to permit adequate assessment of reliability for that group.

Bilingual Syntax Measure-Spanish. The BSM-S (for Spanish-dominant children) is considered conditionally reliable. The fall and spring alphas were acceptable; but the fall-to-spring correlation was low (.38), possibly due to slight scoring changes made during the spring testing period.

WPPSI Block Design. This test, intended to be given at all sites in the fall of Years II and III of the evaluation, was given to the children in the Michigan site this spring since those children were not tested last fall. At all time points, the test is judged to be reliable, based on the alpha coefficients for internal consistency. No fall-spring correlation could be calculated since no children received the test at both time points.

Table 2

Estimates of Reliability of the Child Measures, Based on Cronbach's Alpha (Internal Consistency)[1] For Fall 1975 and Spring 1976--Head Start Children

		Cronbach's Alpha			
		English-Dominant Children		Spanish-Dominant Children	
Child Measures	(Source)	Fall 1975	Spring 1976	Fall 1975	Spring 1976
COGNITIVE-LANGUAGE					
Bilingual Syntax Measure-English		.82 (691)[2]	.88 (430)	.93 (17)	.93 (10)
Bilingual Syntax Measure-Spanish		.88 (13)		.96 (85)	.76 (70)
Block Design	(WPPSI)	.75 (724)	.78 (80)	.80 (87)	
Conceptual Grouping	(MSCA)	.63 (721)	.61 (458)	.63 (87)	.71 (68)
Say and Tell	(CIRCUS)	.47 (720)	.31 (440)	.47 (86)	.40 (68)
Verbal Fluency	(MSCA)	.75 (726)	.74 (458)	.72 (87)	.71 (68)
Verbal Memory-1	(MSCA)	.64 (724)	.73 (435)	.67 (87)	
Verbal Memory-3	(MSCA)	.85 (725)	.83 (434)	.74 (87)	
PSYCHOMOTOR					
Arm Coordination	(MSCA)	.54 (738)	.62 (457)	.58 (87)	.76 (67)
Draw-A-Child	(MSCA)	.82 (737)	.74 (456)	.81 (87)	.67 (67)
Leg Coordination	(MSCA)	.56 (733)	.49 (424)	.53 (87)	.34 (66)
SOCIAL-EMOTIONAL					
POCL	(High/Scope)	.90 (719)	.93 (462)	.87 (87)	.94 (70)
Child Rating Scale-1: "Poise"	(High/Scope)	.82 (205)	.91 (153)		
Child Rating Scale-2: "Agressiveness"		.87 (205)	.75 (153)		
Child Rating Scale-3: "Determination"		.35 (205)	.82 (153)		
Child Rating Scale-4: "Self-sufficiency"		.69 (205)	.85 (153)		
Child Rating Scale-5: "Flexibility"		.75 (205)	.69 (153)		
SITE SPECIFIC MEASURES[3]					
Opposite Analogies (Colorado and Utah)	(MSCA)	.76/.64 (71) (69)	.69 (52)		
Do You Know..? (Florida and West Virginia)	(CIRCUS)	.42/.67 (34) (58)	.72 (46)		

---

[1] The samples consisted of all PDC and Comparison Head Start children across all sites. Alphas were calculated separately for English- and Spanish-speaking children.

[2] Numbers in parenthesis are the sample size on which the Cronbach's alpha is based.

[3] Fall alphas are reported by site; they were not calculated overall.

Table 3

Stability of Child Measures: Correlation of Fall Scores with Spring Scores

	Correlation of Fall 1975 Score with Spring 1976 Score			
	English-Dominant Head Start Children		Spanish-Dominant Head Start Children	
Child Measure	N	r	N	r
BSM-English	331	.68		
BSM-Spanish[1]			68	.38
Conceptual Grouping[1]	92	.49		
Verbal Fluency[1]	362	.45	68	.36
Verbal Memory-1	328	.38		
Verbal Memory-3	342	.43		
Arm Coordination[1]	41	.72	11	.65
Draw-A-Child	361	.51	67	.45
Leg Coordination[1]	27	.45		
POCL	355	.48	70	.39
PIPS[2]	120	.49		

NOTE

All correlations are significant at the $p < .05$ level.

[1] Administration and/or scoring changed from fall to spring, resulting in some reduced sample sizes and possibly reduced correlations.

[2] The PIPS was administered at fewer sites in the fall.

Conceptual Grouping. This test is unacceptable in terms of its reliability for both English- and Spanish-dominant children, due to low alphas (in spite of an acceptable fall-spring correlation).

Say and Tell. Say and Tell, a CIRCUS subtest, obtained unacceptably low alphas and is judged to be unreliable for both samples.

Verbal Fluency. Verbal Fluency is judged to be conditionally reliable, based on acceptable alphas for both language groups and an acceptable fall-spring correlation for the English sample; but the fall-spring correlation was low for the Spanish sample. The correlations were probably somewhat reduced because of a change in item order from fall to spring which was intended to help children distinguish between the animal and toy categories.

Verbal Memory. Parts 1 and 3 of this test were administered, and both have acceptable levels of internal consistency although the fall-to-spring correlation for Verbal Memory-3 was low. The test was not administered in the spring to Spanish-dominant children. Both parts are reliable for the English sample.

Arm Coordination. Some changes in the administration of this test appear to have improved the alphas for the spring testing for both language groups; and since the fall-to-spring correlations were also acceptable, Arm Coordination is judged reliable.

Draw-A-Child. Since the alphas and the fall-spring correlations were all acceptable, Draw-A-Child is considered reliable for both language groups.

Leg Coordination. This test is judged unreliable due to the lack of internal consistency.

Pupil Observation Checklist.[1] The POCL attained a high degree of internal consistency for both language groups at both time points, and the fall-spring correlations were acceptable but lower than anticipated. This measure is judged to be reliable.

Preschool Interpersonal Problem-Solving Test. The scoring of the PIPS does not result in item scores that can be used to calculate Cronbach's alpha. But based on an acceptable fall-spring correlation, the PIPS is judged reliable.

---

[1] The subscales for the POCL and the Child Rating Scale are discussed in Appendix H.

PDC Child Rating Scale. All of the spring factors on the Child Rating Scale are judged to be reliable, based on internal consistency. No fall-spring correlations were calculated since the fall and spring factor scores do not match completely in terms of items included in each. The rating scale was used in only six sites this spring, none of them with Spanish-dominant children, so the reliability estimates apply only to the English-dominant group.

PDC Child Interview (Faces). Testers' reports from the field made it clear that the Faces procedure was so confusing for Head Start children that the results would not be useable. Thus, the standard psychometric analyses were not carried out for this measure. Instead, the data will be reserved for later study that may lead to revision of the instrument.

Validity of the Instruments for Head Start Children

The validation procedure (described more fully in the Methods section of this report) involved determining the expected logical relationship of each reliable measure to others, then comparing these expectations with the relationships that actually appeared in the data. Table 4 displays the matrix of expected relationships. The more consistent the expected and the actual relationships for any measure, the stronger the basis for assuming measurement validity. Under this procedure, degree of validity is expressed in terms of the mean deviation of expected from actual relationships. For example, if it was expected that the correlation of Measure A with Measure B would fall in the range defined as "Medium" but it actually fell in the "low" range, this represented a deviation of 1 unit from the expectation. Deviations were summed and averaged to produce a validity index for each measure.

Figure 3 presents profiles of the indexes so obtained for each measure across testing samples. As in the interpretation of data analyzed in fall 1975, a mean deviation of 1 or less was accepted as evidence of validity (deviations could theoretically range from 0 to 6). According to the criterion, all the instruments examined are acceptably valid for Head Start children with the possible exception of the English version of the BSM when administered to Spanish-dominant children. This finding and validity findings for older children are discussed at greater length in this report.

Table 4

Hypothesized Correlation Matrix for Spring Data, 1976

Scale:

0: -.1 to .1

Low: .1 to .3

Med: .3 to .5

High: .5+

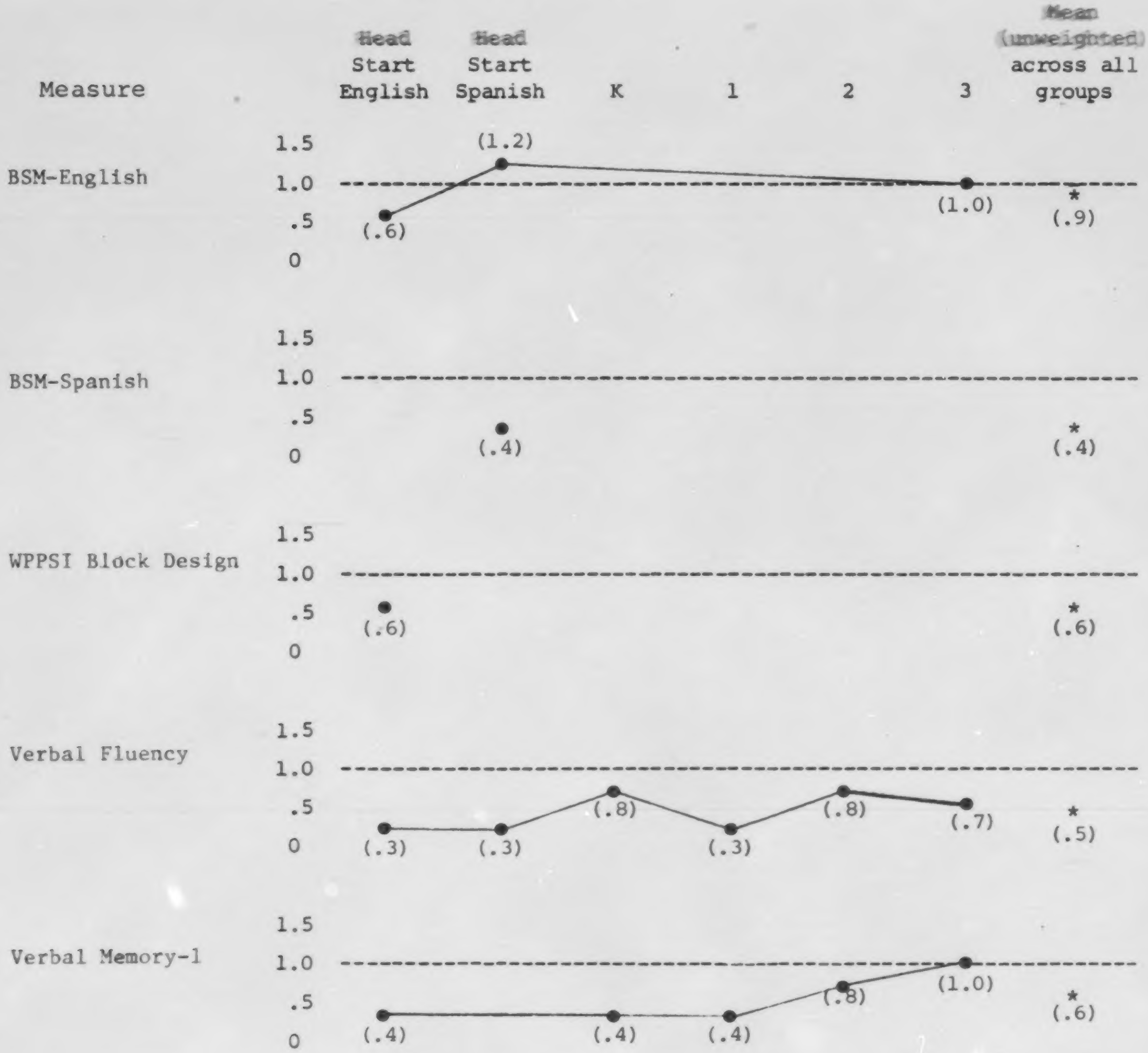
U: Undetermined

	CHILD MEASURES	Bilingual Syntax Measure	Block Design (WPPSI)	Conceptual Grouping	Say and Tell	Verbal Fluency	Verbal Memory-1	Verbal Memory-3	Arm Coordination	Block Building	Draw-A-Child	Leg Coordination	PIPS-Solutions	PIPS-Locus of Control	Pupil Observation Checklist
COGNITIVE-LANGUAGE	BSM	1.00													
	Block Design (WPPSI)	Med	1.00												
	Conceptual Grouping	Low	Med	1.00											
	Say and Tell	Med	Low	Low	1.00										
	Verbal Fluency	Med	Low	Low	High	1.00									
	Verbal Memory-1	Med	Low	Low	Med	Med	1.00								
	Verbal Memory-3	Med	Low	Low	Med	Med	Med	1.00							
PSYCHOMOTOR	Arm Coordination	0	0	0	0	0	0	0	1.00						
	Block Building	Low	Med	Med	Low	Low	Low	Low	Low	1.00					
	Draw-A-Child	Low	Med	Med	Low	Low	Low	Low	Low	Med	1.00				
	Leg Coordination	0	0	0	0	0	0	0	High	Low	Low	1.00			
SOCIAL-EMOTIONAL	PIPS-Solutions	Low	Low	Low	Low	Low	Low	Low	0	Low	Low	0	1.00		
	PIPS-Locus of Control	U	U	U	U	U	U	U	U	U	U	U	Low	1.00	
	POCL	Med	Med	Med	Med	Med	Med	Med	Low	Low	Med	Low	Med	Med	1.00

All non-zero correlations are positive.

Figure 3

Validity Profiles for Measures Judged Tentatively Reliable, Spring 1976 Data



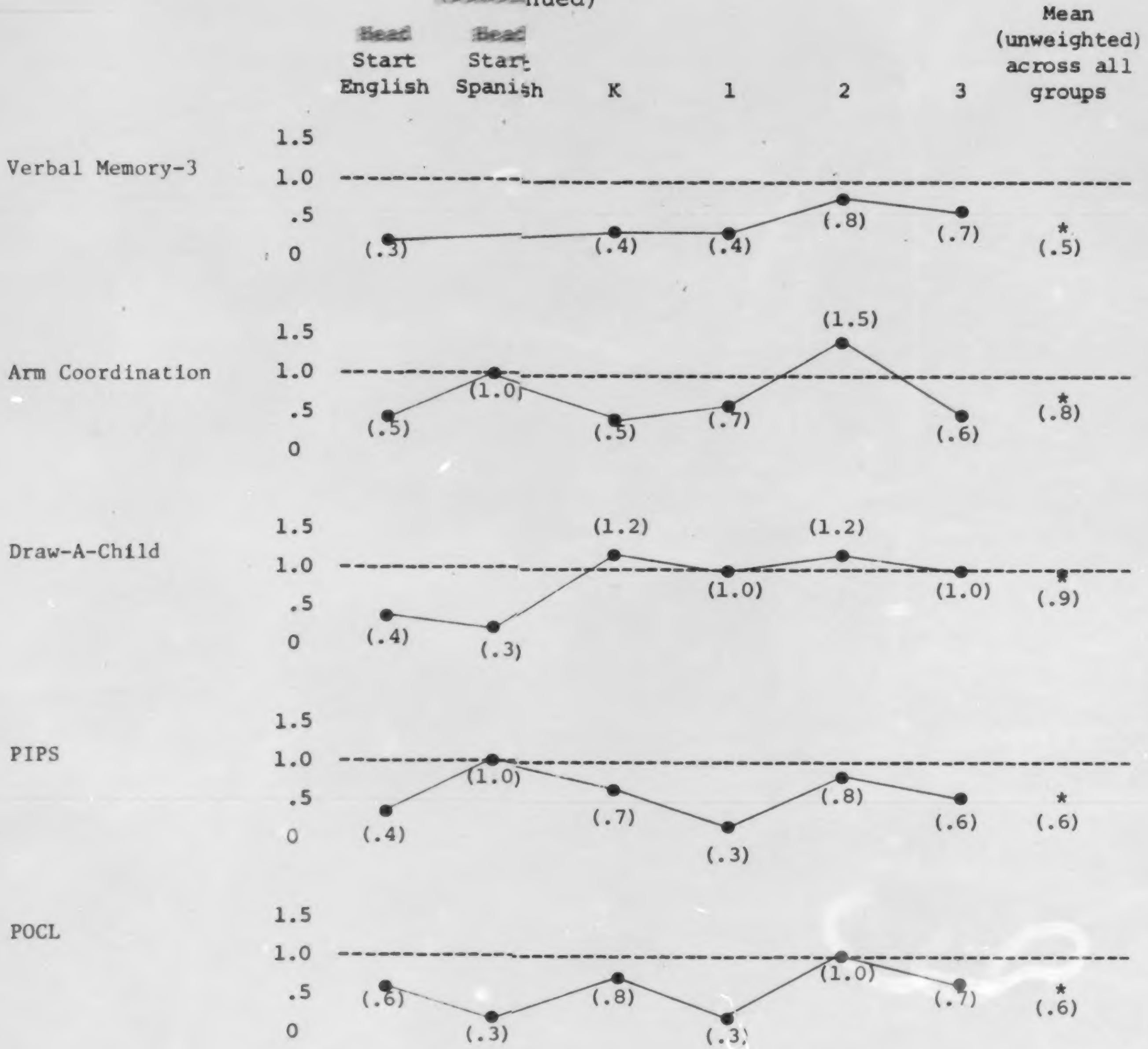
Note

The points plotted represent the mean deviation of the measure from expected relationships with other measures. The broken line represents the level above which deviations are considered excessive. (The range of theoretically possible deviations extends from 0 to 6.)

continued:

Figure 3 (continued)

Validity Profiles for Measures Judged Tentatively Reliable, Spring 1976 Data (continued)



Sensitivity of Child Measures to Change

Since the impact evaluation will be using child measures to detect change that can be attributed to program differences, it is necessary to determine the extent to which the child measures included in the battery are actually sensitive to change. The findings reported here should be considered only estimates of the sensitivity that the measures are likely to show in the future, since (1) the time period between testings was shorter than it will be in the future, (2) it was primarily PDC children who were included in the testing and analysis, and (3) some of the measures underwent slight revisions between the fall and spring testing periods.

As described under Data Analysis Procedures in Chapter II, the sensitivity-to-change analyses were of three types. First, the correlation of each measure, in the fall and in the spring, with child age at the time of testing was calculated to determine the age-relatedness of the measures (a substantial relationship was expected to be found for most measures). Next, a paired (or correlated) t test was calculated for the difference between the fall mean score and the spring mean score on each measure to ascertain if the scores increased significantly from fall to spring. Finally, a regression procedure was used to determine whether the observed spring mean on a measure was equal to or greater than the expected, or predicted, spring mean, which would show that children gained at least as much as they were expected to gain over the given time interval. All three analyses were examined to assess sensitivity to change. The results of the analyses were more critical for some measures than for others. For example, scores on the Child Rating Scale and the POCL were expected to be less related to age than were scores on the other measures since the ratings were being made relative to other children their age. And while scores on these ratings were expected to change from fall to spring, they were not expected to change in a consistent up-or-down direction.

Correlations with age. The correlations of each of the measures with child age at the time of testing are shown in Table 5. The correlations tend to be significant but low, generally around .15 to .30. They were not, in fact, expected to be much higher, since the measures would not be useful for a program evaluation if they were related mainly to age rather than to differential experiences. The only consistent problem indicated by the correlations is with Draw-A-Child for the Spanish-dominant group. There was, however a significant gain from fall to spring (see Table 6) for these children. In addition, consideration is being given to a more complex scoring

Table 5

Sensitivity of Child Measures to Change: Correlations of Fall 1975 Scores and Spring 1976 Scores With Age at Time of Testing

	Correlation of fall score with fall age		Correlation of spring score with spring age	
	English	Spanish	English	Spanish
BSM-E	.25** (328)		.17** (406)	
BSM-S		.33** (68)		.28** (70)
Conceptual Grouping	.22** (92)		.21** (442)	.25** (68)
Verbal Fluency	.20** (356)	.23* (70)	.20** (435)	.17 (68)
Verbal Memory-1	.12** (339)	.25* (69)	.08* (420)	
Verbal Memory-3	.22** (355)	.22* (70)	.13** (420)	
WPPSI Block Design			-.10 (79)	
Arm Coordination			.16** (434)	.54** (67)
Leg Coordination			.19** (410)	.38** (66)
Draw-A-Child	.29** (356)	.13 (70)	.18** (433)	.12 (67)
PIPS	.19** (123)		.15** (423)	.18 (68)
POCL (total) fall	.22** (347)	.00 (70)		
POCL-1 spring			.17** (435)	.28* (70)
POCL-2 spring			.15** (435)	.23* (70)
Do You Know			.21 (44)	
Opposite Analogies			.42** (48)	
CRS-F1: lead	.14 (95)			
CRS-F2: nice	.14 (95)			
CRS-F3: work	.30** (95)			
CRS-F4: mime	.05 (95)			
CRS-S1: poise			.03 (145)	
CRS-S2: agressiveness			.03 (153)	
CRS-S3: determination			.13 (152)	
CRS-S4: self-sufficiency			.09 (155)	
CRS-S5: flexibility			.25** (155)	

(N is in parenthesis)

*significant at p<.05, directional probability

**significant at p<.01, directional probability

Table 6

Sensitivity of Child Measures to Change: Difference Between Fall 1975 and Spring 1976 Scores, as Measured By Pairwise t Tests

	English-Dominant Head Start Children				Spanish-Dominant Head Start Children			
	N	Fall Mean	Spring Mean	Pairwise t statistic	*N*	Fall Mean	Spring Mean	Pairwise t statistic
BSM-E	*326*	9.4	10.7	7.97**	*5*	.2	1.8	1.0
BSM-S	*4*	2.0	5.3	1.27	*68*	10.1	12.6	4.56**
Conceptual Grouping	*92*	5.4	9.3	11.3**				
Verbal Fluency	*355*	6.1	9.4	11.96**	*68*	4.4	6.9	5.13**
Verbal Memory-1	*328*	10.0	11.1	5.48**				
Verbal Memory-3	*342*	2.4	4.1	10.28**				
Arm Coordination	*39*	5.0	6.0	2.17**	*11*	2.7	3.8	2.06*
Draw-A-Child	*354*	4.2	5.9	12.99**	*67*	5.3	6.1	2.96**
Leg Coordination	*27*	10.0	9.3	-1.18				
PIPS	*120*	2.0	2.9	5.55**	*68*		2.3	

*significant with p<.05, directional probability

**significant with p<.01, directional probability system for this measure, as indicated in the section "Suitability of the Measures for Older Children," and this change might improve the score-age correlation. The other noted problems are with the spring scores for Verbal Fluency and PIPS, again for the Spanish-dominant group; here the correlations are not noticeably different from those for the English-dominant group, but are not significant due to the smaller sample size.

WPPSI Block Design shows a small negative correlation with age. This correlation was not expected to be negative, but neither was it expected to be significantly positive because of the restricted range of children's ages. The test was chosen as a maturation index, to be used as a covariate for other analyses, and to be administered only in the fall of each year. It was administered to 79 children this spring in the Michigan site because those children had not been tested last fall.

Fall-to-spring change: *t* tests. Table 6 shows the fall and spring means on each measure for each language group and the *t* statistic for the differences between means. The Bilingual Syntax Measure-English showed a non-significant increase for Spanish-dominant children, and the BSM-S showed a non-significant increase for English-dominant children. The sample sizes were, however, extremely small; so the measure may or may not be useful for children who are not proficient in the language tested. All other child measures showed a significant fall-to-spring increase, except Leg Coordination, which is being eliminated from the evaluation because of low reliability.

Fall-to-spring change: regression analysis. For four of the child measures (BSM-E, Draw-A-Child, Verbal Fluency and Verbal Memory-3), R^2 values are comparable for fall and spring regression equations which relate child status variables (age, sex, ethnicity, preschool experience, and siblings) to child score on the measure. For each of these four measures, therefore, an expected spring score was calculated based on fall score, status on background variables, and spring age. This expected spring mean is shown in Table 7, along with the actual spring mean and the *t* statistic for the difference between the means. For Draw-A-Child, Verbal Fluency, and the Verbal Memory Scales, the actual spring mean was greater than the expected spring mean; the children gained more on the measures than was expected as a function of their increase in age. For all four of the measures, more than half of the children obtained an actual spring score that was equal to or greater than their expected spring score. These results imply that the tests are sensitive to change that is due to educational experience in addition to experience that is simply a function of increased age. The fact that this

Table 7

Sensitivity of Child Measures to Change: Difference Between Actual and Projected Spring 1976 Scores (Projections Based on Fall 1976 Scores and Change in Age; English-dominant children)

Child Measure[1]	N	Projected Spring Mean	Actual Spring Mean	Pairwise t statistic	Percent of children whose actual spring score was ≥ projected spring score
BSM-E	261	10.7	10.5	-1.36	54.5
Draw-A-Child	288	5.3	5.8	3.59**	61.8
Verbal Fluency	289	7.8	9.2	4.6**	58.1
Verbal Memory-3	279	3.4	3.9	3.11**	57.0

**significant with $p<.01$, directional probability

[1] This analysis was conducted only on those child measures for which the R^2 values for fall and spring prediction equations were comparable (see text).

analysis was not performed for the other measures does not mean that those tests are not sensitive to such change; they were excluded because the statistical assumptions underlying the analysis did not appear tenable for those particular tests.

Summary of sensitivity to change. Based on the results of the three analyses discussed above, it is concluded that all the measures that were expected to detect change over the fall-spring testing interval are in fact sensitive to change. Therefore the subset of measures that have been judged reliable and valid appear suitable for use in the longitudinal evaluation, at least for children within the developmental range of the present sample.

Suitability of the Instruments for Use in the Higher Grades

Part of the task of assessing the usefulness of the battery of child measures for the PDC evaluation has been to determine how suitable they are for use with children in kindergarten through grade 3 as well as for use with Head Start children. As part of the cross-sectional design in the Georgia site, approximately 25 children per grade were tested (kindergarten through grade 3), and in Maryland 30 third-graders were tested. This information was used to determine how suitable the instruments can be expected to be for future use at those grade levels.

Conclusions about the suitability of the child measures for use in kindergarten through grade 3 were based on four factors, which were considered at each age level: response distributions on the items of each measure, mean scores on each measure, reliability (internal consistency) of the measures, and validity of the measures.

The response distributions for each age level (Head Start through grade 3) are contained in Appendix D. An item was considered not to be useful at a given grade level if more than 75% of the children at that level received the maximum score for the item. Table 8 shows the number of items for each measure that were thus judged not to be useful at each age level.

The mean scores on the child measures at each grade level are presented in Table 9. Since the POCL and Child Rating Scale (CRS) are rating instruments, wherein children are judged in relation to their peers, the mean scores do not increase or decrease systematically across the grade levels. If the mean score at any age was greater than 80% of the total possible score for the test, the measure was considered unsuitable for children of that age.

The estimates of reliability for the measures at each grade level are presented in Table 10, and the validity profiles have been shown in Figure 3. The criteria for reliability and validity for children in higher grades were essentially the same as those for Head Start.

Based on the above criteria for determining suitability of the instruments, most of the measures appear to be useful through grade 3, either in their present forms or with modification. Each child measure is discussed below.

Table 8

Suitability of Measure for Higher Grades: Number of Non-useful Items at Each Grade Level, Spring 1976 Data

	Number of items on which more than 75% of children tested received the highest score.					Total number of items on test
Child Measure	Head Start	K	Grade 1	Grade 2	Grade 3	
BSM-English	5				12	18
BSM-Spanish	10					18
Conceptual Grouping	0	1	1	2	2	9
Verbal Fluency	0	0	0	0	0	4
Verbal Memory-1	1	2	2	2	4	4
Verbal Memory-3	0	0	0	1	0	11
Arm Coordination	0	0	0	0	2	6
Draw-A-Child	3	6	9	9	9	10
Leg Coordination	2	3			5	6

Table 9
Mean Scores on Child Measures for Each Grade Level, Spring 1976 Data

	Head Start Total English Sample	Kdg. Georgia	Grade 1 Georgia	Grade 2 Georgia	Grade 3 Georgia	Grade 3 Maryland	Maximum Possible Score
N[1]	424-462	26	27	27	24	30	
Average Age in Months	59.4	71.0	84.1	96.7	108.1	108.3	
CHILD MEASURES							
BSM-English	10.2					14.1	18
Conceptual Grouping	6.7	7.9	9.6	10.2	10.6	10.5	12
Verbal Fluency	9.1	14.1	16.4	20.8	21.8	22.5	36
Verbal Memory-1	10.8	11.8	12.6	13.0	13.4	13.6	14
Verbal Memory-3	4.0	6.8	6.3	7.3	8.2	8.0	11
Arm Coordination	5.0	8.5	11.0	14.6	18.5	18.5	28
Draw-A-Child	5.9	7.7	8.7	9.0	9.4	9.5	10
Leg Coordination	8.8	10.0				10.9	13
PIPS	2.8	3.8	3.3	4.9	4.4	5.5	7
POCL-1	32.8	32.0	31.5	33.7	32.6	40.8	56
POCL-2	12.8	12.4	11.6	12.4	11.9	14.2	21
CRS-1 (N=159)	35.0	35.7	34.4	32.2	35.9		55
CRS-2 (N=159)	8.4	7.2	7.9	7.2	6.0		15
CRS-3 (N=159)	-25.7	-26.2	-24.3	-25.4	-21.7		-45
CRS-4 (N=159)	20.7	20.2	20.6	18.9	20.7		30
CRS-5 (N=159)	15.2	15.7	16.3	14.1	16.1		25

---

[1] Sample sizes are approximate and vary most greatly for the Head Start sample but only slightly for other grades.

Table 10

Estimates of Reliability of the Child Measures Based on Cronbach's Alpha (Internal Consistency)[1] For Fall 1975 and Spring 1976--Elementary Grades in Georgia and Maryland[2]

		Cronbach's Alpha							
Child Measures		K Fall	K Spring	1 Fall	1 Spring	2 Fall	2 Spring	3 Fall	3 Spring
COGNITIVE-LANGUAGE									
Bilingual Syntax Measure-English									.58 (30)[3]
Conceptual Grouping	(MSCA)	.67 (30)	.60 (26)	.48 (28)	.51 (27)	.41 (31)	.28 (27)	.30 (29)	.35 (54)
Say and Tell	(CIRCUS)	.49 (32)							.44 (30)
Verbal Fluency	(MSCA)	.77 (30)	.39 (26)	.75 (28)	.71 (27)	.26 (31)	.79 (27)	.67 (29)	.69 (54)
Verbal Memory-1	(MSCA)	.81 (30)	.46 (24)	.55 (28)	.51 (27)	-.01 (31)	.27 (27)	.07 (29)	.19 (54)
Verbal Memory-3	(MSCA)	.85 (30)	.76 (24)	.86 (28)	.78 (27)	.80 (31)	.66 (27)	.74 (29)	.50 (54)
PSYCHOMOTOR									
Arm Coordination	(MSCA)	.69 (30)	.78 (26)	.64 (28)	.68 (27)	.70 (31)	.76 (27)	.69 (29)	.65 (54)
Draw-A-Child	(MSCA)	.70 (30)	.65 (26)	.61 (28)	.61 (27)	.58 (31)	.44 (27)	.43 (29)	.21 (53)
Leg Coordination	(MSCA)	.45 (30)	.42 (24)						
SOCIAL-EMOTIONAL									
POCL	(High/Scope)	.95 (30)	.54 (26)	.93 (28)	.86 (27)	.91 (31)	.91 (27)	.83 (29)	.97 (54)
Child Rating Scale-1: "Poise"	(High/Scope)		.96 (26)		.85 (24)		.93 (26)		.85 (24)
Child Rating Scale-2: "Agressiveness"			.56 (26)		.79 (24)		.78 (26)		.66 (24)
Child Rating Scale-3: "Determination"			.89 (26)		.80 (24)		.74 (26)		.77 (24)
Child Rating Scale-4: "Self-sufficiency"			.97 (26)		.97 (24)		.77 (26)		.93 (24)
Child Rating Scale-5: "Flexibility"			.33 (26)		.87 (24)		.71 (26)		.78 (24)

---

[1] The samples consisted of random selections of elementary children in the Georgia PDC school.

[2] Maryland data enter only into the spring grade 3 column.

[3] Numbers in parenthesis are the sample sizes.

BSM-English and BSM-Spanish. These measures will probably be suitable for the evaluation through grade 3. Intermediate grades were not tested on these measures, so not enough information is available to make a definite decision regarding them. The mean scores and validity profiles are generally acceptable, although the Spanish-language version shows greater deviation from the validity criterion than might be desired. The number of non-useful items and the third-grade internal consistency (alpha = .58) make the use of these measures in the upper grades problematic.

Conceptual Grouping. A decision has already been made to drop this measure from the evaluation based on results at the Head Start level, but it seemed possible that it might be useful for older children. However, since the reliabilities are poor in grades 1 through 3, and since the mean scores for grades 2 and 3 are unacceptably close to the maximum score, Conceptual Grouping will not be considered for use with the upper grades.

Verbal Fluency. Based on all four considerations, Verbal Fluency is judged to be suitable for use at all grade levels, Head Start through grade 3.

Verbal Memory-1. This test, very near ceiling at Head Start, is unusable beyond kindergarten because most of the children receive the maximum score. The task is too simple for the older children, and this test (or one like it) will be used beyond Head Start only if more difficult items can be constructed.

Verbal Memory-3. Based on all four considerations, Verbal Memory-3 is judged to be suitable for use at all grade levels.

Arm Coordination. The only indication of difficulty with this measure is in the validity profile, where the deviation from expected correlation levels is rather large at grade 2. However, the deviations appear to peak at grade two for all of the measures, so this sample may not accurately represent the general population. Arm Coordination is therefore judged to be suitable for all grade levels.

Draw-A-Child. As with Verbal Memory-1, scores on this measure approach the maximum score by grade 1, making it unusable in its present form beyond kindergarten. The problem with this test, however, appears to be in the scoring rather than in the simplicity of the task. We have used a less complex scoring procedure than that recommended by the test author. We would now recommend returning to the more complex procedure or to a

Goodenough-Harris type of scoring in order to make use of the available information that the drawings of older children can provide. With a change in scoring procedures, Draw-A-Child is expected to be suitable for use with all age levels.

Leg Coordination. A decision was made to drop this measure from the evaluation based on results at the Head Start level. Data from the higher grades support this decision.

PIPS. Based on all four considerations, the PIPS is judged to be suitable for use at all grade levels.

POCL and Child Rating Scale. Both rating scales are judged to be suitable for use at all grade levels. Mean scores, alphas, and validity appear to be acceptable. The response distributions are acceptable, but tend to be unusually centrally-distributed for the POCL for grades 1 through 3, which may be attributable to the particular testers who assign the ratings (this will be investigated in the future).

Factor Structure of the Battery

To investigate relationships among the child measures, the measures or subscales found reliable were factor analyzed[1], producing the results shown in Tables 11 and 12. Separate factor analyses were performed for the English- and Spanish-dominant groups. The two analyses are not directly comparable, however, since not all of the same scales were used; Spanish-language children were not tested on Verbal Memory nor was the Child Rating Scale administered at their sites in spring 1976.

Results for the English-dominant sample. The subscales of the Child Rating Scale loaded on factors that do not appear to be primary factors for any of the other measures in the battery, suggesting that these scales predominantly represent other dimensions. The two scales of the Pupil Observation Checklist, however, loaded highest on a factor that appears to be the principal factor for three of the tests: Verbal Fluency, Verbal Memory-3, and the Preschool Interpersonal Problem-Solving Task (PIPS). It is interesting to note that Arm Coordination seems to tap a dimension that is largely independent of those represented by the other measures--presumably a psychomotor dimension--and that Draw-A-Child, as expected, is more highly related to cognitive tests than to Arm Coordination.

Results for the Spanish-dominant sample. The first factor emerging from this analysis closely resembles the second factor found in the preceding analysis: for Spanish-dominant children, Verbal Fluency and the two POCL scales loaded principally on the same factor, a factor on which the PIPS also loaded substantially. The second factor, though, is the primary factor for four of the five tests, and, surprisingly, Arm Coordination has the highest loading on this factor of any measure. It is not readily apparent why the parallel with the English-dominant sample fails here--investigation will continue in future analyses.

---

[1] Principal components solution, varimax rotation.

Table 11

Factor Analysis[1] of Scores on Child Measures, English-Dominant Head Start Children, Spring 1976 Data

N = 118

Child Measure	Factor Loadings of Child Measures (highest loading italicized)				
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
BSM-English	.30	-.03	*.77*	-.24	.08
Verbal Fluency	.21	*.49*	.45	.10	.08
Verbal Memory-1	-.13	.35	*.66*	-.08	-.28
Verbal Memory-3	.00	*.64*	.20	-.20	.15
Arm Coordination	-.04	-.12	-.03	*.92*	.04
Draw-A-Child	.15	.13	*.61*	.29	-.22
PIPS	.16	*.73*	.10	-.02	-.06
POCL					
1. "Task Orientation"	.14	*.74*	.23	.08	-.34
2. "Extroversion"	.23	*.69*	-.17	-.07	.29
Child Rating Scale					
1. "Poise"	*.81*	.24	.05	.03	.19
2. "Aggressiveness"	-.02	.09	-.07	.09	*.87*
3. "Determination"	.50	.01	.17	.16	*-.66*
4. "Self-sufficiency"	*.83*	.15	.03	-.10	-.27
5. "Flexibility"	*.80*	.12	.31	-.01	-.17

[1] Principal components solution, varimax rotation.

Table 12

Factor Analysis[1] of Scores on Child Measures, Spanish-Dominant Head Start Children, Spring 1976 Data

N = 66

Child Measure[2]	Factor Loadings of Child Measures (higher loading italicized)	
	Factor 1	Factor 2
BSM-Spanish	.31	*.53*
Verbal Fluency	*.43*	.40
Arm Coordination	.13	*.72*
Draw-A-Child	-.05	*.67*
PIPS	.45	*.53*
POCL		
1. "Task Orientation"	*.90*	.22
2. "Extroversion"	*.87*	-.02

---

[1] Principal components solution, varimax rotation.

[2] Children in the Spanish-dominant sample did not receive the Verbal Memory test and were not rated on the Child Rating Scale in Spring 1976.

Relationship of the Tests to Social Competence

Rationale for selection of "social competence" criteria. Social competence, as it is understood in this study, means everyday effectiveness; it is probably close to the familiar ideas of "savvy" and "know-how"--ability to handle oneself in a variety of situations. But although the term "ability" implies a singular trait, there can be little doubt that getting along in life calls for multiple skills; they can be categorized in many different ways--social, physical and intellectual, for example--and these categories, too, can be subdivided. However, division of social competence into component traits can seem to imply that each component is empirically independent of the others just because it is conceptually independent, i.e., that because social, physical, and intellectual abilities can be thought of separately, their occurrence in life is unconnected. But it is not difficult to contradict this implication by referring to familiar personality types in which these traits are interdependent (for example, highly verbal children are often highly sociable children, and vice-versa). It may be misleading, then, to consider social competence to be simply the sum of all conceivable abilities (Anderson & Messick[1] call this the "bag of virtues" view); instead, it may be more reasonable to view social competence as a special conjunction of abilities, a practical alignment of the various talents that it takes to manage successfully in day-to-day life.

Regarded in this way, then, social competence consists of elements that can be identified separately, but that function interdependently. An analogy (admittedly exaggerated) can be drawn with the human body: the various organs each have their specialized place and purpose, but all must function together to function at all. This view of social competence is strictly hypothetical at the moment, since there are a number of alternative views and little basis in research for any of them. But it is a working hypothesis--one that will be tested as the evaluation of PDC progresses.

Extraction of "social competence" factor scores. In the analysis of spring data, the social competence of children in the sample was represented by measures extracted from the PDC Child Rating Scale and the Pupil Observation Checklist (hereafter abbreviated as CRS and POCL), instruments on which classroom teachers and testers respectively rated the child on a variety of dimensions, e.g., "Has the desire to master all kinds of skills;" "Is Active. . . . . . .Passive." These two instruments were selected as proxy "social competence" criteria because (1) they appear to tap a wider range of traits than any of the tests, (2) the traits represented in them closely resemble practical notions of social competence, and (3) it is possible that, if social competence does involve alignment of interdependent traits as hypothesized, teachers and testers are intuitively able to take the degree of alignment into account in assigning ratings, so that the ratings correspond not simply to the child's possession of molecular characteristics, but to his or her ability to "get it together."

---

[1] Anderson, S. & Messick, S. Social competency in young children. Developmental Psychology, 1974, 10, 282-293.

The first step in the analysis was to pool and factor analyze[1] the 39 items of the CRS and the 11 items of the POCL to identify clusters of homogeneous items. The CRS was administered at only six sites, thus only a subset (142 children) of the spring testing sample entered into the factor analysis. Since the CRS was not administered in either the Texas or the California site, no analysis was possible for Spanish-dominant children.

Table 13 describes the factors that emerged from the analysis; each item is listed under the factor on which it loaded highest. (The labels assigned to the factors represent what appear to be the common features of the items contributing most heavily to each.) Next, factor scores were created by multiplying each child's scores on each item by a factor coefficient and summing the products for the respective factors. The nine resulting factor scores were taken as representing "social competence" for purposes of this analysis. As a check on the assumption that these factor scores measured central dimensions of performance, a second factor analysis was carried out on the nine factor scores pooled with children's spring scores on seven major tests in the battery. It was expected that the "social competence" criteria, if they were indeed such, would load on the general factors that emerged from the analysis, along with the tests (assuming that these too were relevant to social competence).

Only one of the "social competence" criteria loaded on the first factor, along with four of the seven tests; but on each of the factors on which the three other tests loaded highest, one of the criteria also loaded highest. These

---

[1] Principal components solution, varimax rotation.

Table 13

Proxy Social Competence (SC) Criteria:
Item Clusters Produced by Factor Analysis of the PDC Child Rating Scale (CRS) and the Pupil Observation Checklist (POCL)

Factor Loading	
	SC-1: "Cooperativeness"*
-.82	13. Shows physical dislike or hostility to others.
.77	8. Accepts or abides by school, classroom rules.
-.76	2. Uses physical force to try to control others.
-.73	11. Shows verbal dislike or hostility to others.
.70	9. Cooperates and shares with others.
-.53	6. Competes with others for toys, attention, achievement.
.49	17. Gets the attention of adults appropriately.
.48	18. Gets the attention of peers appropriately.
	SC-2: "Social Ease"*
.78	24. Talks freely to adults.
.77	23. Talks freely to children.
.70	15. Shows awareness of and pride in own culture.
.68	1. Uses words or wits to try to influence others.
.65	10. Shows verbal affection to others.
.61	3. Succeeds in influencing or controlling others.
.60	14. Shows self-confidence.
.49	7. Asserts his/her rights to fair treatment.
	SC-3: "Task Orientation"**
.89	Needs urging--Quick to Respond
.86	Gives up easily--Keeps trying
.83	Prefers easy tasks--Attempts difficult tasks
.81	Resistant--Cooperative
.80	Indifferent--Involved
.77	Easily distracted--Attentive
.75	Needs reassurance, praise, encouragement--Realistically self-confident
.71	Nervous--Relaxed
	SC-4: "Achievement Motivation"*
.65	30. See his/her errors in a task and corrects them.
-.64	39. Needs much encouragement or material rewards to attempt tasks.
-.63	27. Is easily distracted when doing a task.
-.63	31. Gives up on tasks before they are finished.
.62	26. Uses alternative strategies if initial problem-solving methods fail.
	SC-4 (continued)
-.58	5. Imitates others or follows them around.
.57	37. Expects to succeed, not afraid of failure.
.56	38. Has a desire to master all kinds of skills.
.55	32. Returns to unfinished tasks after interruption.
-.54	4. Is controlled or influenced by others.
-.47	29. Asks for unnecessary help to do a task.
-.44	35. Chooses tasks that are too easy for him/her.
	SC-5: "Asks for needed help"*
.63	28. Asks for needed help to do a task.
	SC-6: "Self-contentment"*
.77	33. Enjoys tasks he/she chooses.
.64	34. Shows pride in what he/she does or makes.
.52	25. Works or plays well on his/her own.
.48	16. Decides for self what to do, with whom to play.
	SC-7: "Extroversion"**
.77	Quiet--Talkative
.72	Passive--Active
.69	Shy--Sociable
	SC-8: "Chooses hard tasks"*
.60	36. Chooses tasks that are too hard for him/her.
	SC-9: "Maturity"*
.67	22. Recognizes others' feelings, responds appropriately.
.55	21. Shows respect for or tolerance of others' ideas and behavior or looks.
.51	20. Takes on the role of adult during dramatic play.
.51	12. Shows physical affection to others.
.45	19. Attempts to solve social problems with little adult assistance.

---

* Prominent items all from CRS.

**Prominent items all from POCL.

findings offered reasonably convincing support for the assumption that the "social competence" criteria were relevant to general performance, thus the analysis proceeded to the next phase. For this phase, "social competence" variables were also created for the 104 children in the kindergarten-to-grade-3 Georgia comparison sample. Since data from this sample had not entered into the original factor analysis, scale scores, rather than factor scores, were computed by summing the ratings they received on the items assigned to each "social competence" factor.

Statistical relationship of test scores to "social competence" criteria. Having established a basis for tentatively regarding the "social competence" criteria as reasonable proxies for a measure of social competence, it remained to determine the degree to which the tests included in the PDC battery were related to these criteria. The stronger the relationship, the stronger the reason to believe that the traits measured by the tests are genuinely relevant to social competence. To examine the relationship, each of the tests that had been judged reliable and valid (if the validation procedure was applicable) was regressed on the nine "social competence" criteria and on three background variables (ethnicity, age, and sex). The object of the analysis was to determine how fully the test score could be predicted from knowledge of all these variables, and particularly to determine the contribution of "social competence" criteria to this prediction beyond what could be predicted simply from knowledge of the child's status on the background variables. The better the prediction permitted by the criteria, the greater the apparent relevance of the test to social competence. Table 14 displays the findings for each of the tests.

For Head Start children, all of the seven major tests except Arm Coordination were significantly predictable from the "social competence" criteria. Among these, the PIPS showed the strongest relationship to the criteria: 37% of the variance in this test was predictable from knowledge of the "social competence" scores. Since the child's task in the PIPS is to generate solutions to hypothetical interpersonal problems--a task that presumably taps both social and intellectual skills--it is reasonable that this test emerges as the one most closely connected with the "social competence" criteria. The single variable most strongly related to PIPS score is the first POCL factor score, "Task Orientation" (SC-3). Although a large proportion of the variance in Do You Know...? is accounted for by the "social competence" criteria, the small sample size allows little confidence in the stability of this

Table 14

Relationship of "Social Competence" (SC) Criteria and Background Variables to Test Scores of Head Start and Elementary-Level Children

Test	*N*		Percent variance accounted for jointly by "social competence" criteria and background variables		Percent variance accounted for by "social competence" criteria beyond background variables		Individual variables found to be significant predictors ($p \leq .05$) and their partial correlations with test score			
	Head Start	K-3	Head Start	K-3	Head Start	K-3	Head Start Variable	Head Start Partial	K-3 Variable	K-3 Parti
BSM-English	*114*		41***		17**		Ethnicity SC-4 SC-3 SC-8	-.46 .34 .25 .22		
Arm Coordination	*137*	*87*	18*	55***	05	12**	Sex	-.30	Age Sex	.53 -.2
Draw-A-Child	*137*	*87*	29***	36***	20***	09	SC-3 SC-4	.34 .30	Age Ethnicity	.4 -.32
Verbal Fluency	*137*	*87*	29***	54***	26***	14*	SC-3 SC-4	.44 .20	Age SC-7 Ethnicity SC-3	.5 .3 -.32 -.2
Verbal Memory-1	*136*	*86*	22**	19	19**	10	SC-3 SC-4	.35 .23		
Verbal Memory-3	*136*	*86*	30***	40***	27***	25***	SC-3 SC-2 SC-7	.41 .25 .22	SC-6 SC-4 Age SC-2	-.3 .26 .2 .2
PIPS	*136*	*86*	41***	28*	37***	25**	SC-3 SC-7 SC-5 SC-4 SC-2	.50 .29 .27 .19 .18	SC-4 SC-7 SC-9	.29 .2 -.2
Do You Know..?	*24*		77*		68*		Ethnicity SC-4	-.57 .54		
Opposite Analogies	*37*		65**		25		Age SC-4 Ethnicity SC-3	.53 .40 .37 .36		

*Probability of associated F-ratio $\leq .05$.
**Probability of associated F-ratio $\leq .01$.
***Probability of associated F-ratio $\leq .001$.

Key to Factor Scales

SC-2	"Social ease"
SC-3	"Task orientation"
SC-4	"Achievement motivation"
SC-5	"Asks for needed help"
SC-6	"Self-contentment"
SC-7	"Extroversion"
SC-8	"Chooses hard tasks"
SC-9	"Maturity"

figure; similarly, the small size of the sample for which Opposite Analogies data were available probably accounts for the non-significance of the analytic result. For Head Start children, only Sex was a significant predictor of Arm Coordination score--none of the "social competence" criteria, singly or jointly, showed a significant relationship. For elementary-level children, the "social competence" criteria account for 12% of the variance in Arm Coordination; however, although the criteria taken together permit statistically significant prediction, no single one of the criteria reached significance. For children in the elementary-level sample, the "social competence" criteria were not significant predictors of score on Draw-A-Child or Verbal Memory-1. It is probable that this is due to the relatively flat score distributions found for older children on these instruments--the reduced variance results in reduced covariance (with other variables) and hence less predictive precision.

Summary of findings on "social competence." There does appear to be merit in the technique employed here for examining the relationship of child tests to social competence; this approach will be continued and perhaps elaborated in the future. Of the tests examined, Verbal Fluency, Verbal Memory-3, and PIPS all show evidence of relevance to social competence for children from Head Start to grade 3. Verbal Memory-1 and Draw-A-Child, as presently constituted, are apparently related to social competence for children in Head Start but not for older children; in both cases this is probably a consequence of reduced variance in the test scores of older children. For Draw-A-Child it is possible that this can be resolved by substitution of a scoring method that permits greater differentiation among drawings, but the problem for Verbal Memory-1 seems to be in the simplicity of the task itself: repetition of short strings of words. The BSM and Do You Know...? show evidence of relationship to social competence among Head Start children, but whether the relationship would hold for older children cannot be determined, since there were none who had complete data on these tests and on the CRS and POCL variables required for this analysis. Arm Coordination, finally, may be weakly related to social competence as represented here; among children in the K-3 sample, a small proportion of the variance in Arm Coordination scores is predictable from the "social competence" criteria--if the connection is genuine rather than spurious, it is likely that it appeared among older children simply because there is a broader range of performance (greater variance) among these children.

General Analysis of the PDC Classroom Observation System

The PDC Observation System was developed to provide descriptive information regarding the social-emotional competence of children in their classroom settings. The behavior categories that make up the instrument were formed by redefining, and in some cases combining, behavior categories from existing observation instruments that differentiate between socially "competent" and "incompetent" children, and by adding other categories appropriate to the goals of Project Developmental Continuity. The underlying theoretical rationale for selecting these categories is that they can measure a "general attitude of negotiation and reciprocity in dealing with others in our social environment."[1] This attitude is believed to be generalizable across all cultural groups and implies that a child's own needs and goals are valuable, but that the needs and goals of others are equally important and must be taken into account by the child. More specifically, the developing child should learn how to control and influence others with effective strategies that do not violate the rights of others. Accordingly, physical force is considered to be a violation of others' rights, and thus does not indicate an attitude of negotiation and reciprocity. In addition, the child should be reasonably influenced by others, but not totally subservient to or dominated by others. Other social strategies that promote and sustain social interaction such as sharing, helping, requesting and providing resources, and taking turns are also considered important indications of a child's social competence and are represented in the category system.

Fall observation findings. Information garnered from the analysis of fall observation data indicated that observers encountered minimal difficulty while using the instrument in the field, and as much as possible they were able to closely follow standard observation procedures. In a practical sense then, it appeared that classroom observation procedures were a feasible assessment method for a large-scale evaluation effort.

With respect to psychometric characteristics, analysis of the fall observation data revealed that the instrument could detect differences in children's behavior across sites.

---

[1] Bronson, M. Executive competence in preschool children. Paper presented at the meeting of the American Educational Research Association, Washington, D.C., 1975. For a more extensive listing of references to the literature consulted in developing the system, see Interim Report II, Part B (1975).

Further, observations were moderately related to teacher's ratings of the same behavior, attesting, to some degree, to their validity as measures of children's social skills (see Interim Report III, Part A, March, 1976). However, the inter-observer reliabilities of the categories were not determined and the findings indicated low relationships between observation variables and the other instruments in the PDC battery. Thus, information and inferences based on the fall analysis were considered tenative. Nevertheless, this analysis did provide useful information for instrument development that was incorporated into revisions of behavior categories, spring observation training, and data analysis procedures.

Instrument revisions. In the fall, observers indicated that two behaviors, "expression of pride" and "dramatic play," occurred infrequently in the classroom. These reports were substantiated in the later quantitative analysis. Therefore, as proposed in Interim Report III, these categories were eliminated from the instrument. In addition, observers indicated some concern with the complexity of the instrument. In response, it was decided that the observations should center on reciprocal social interactions between the focal child and other persons. Thus, categories that did not directly relate to these interactions were also deleted from the instrument, allowing observers to attend more effectively to children's social interactions. Definitions of other categories which observers identified as ambiguous or unclear were revised and further clarified. These definitions and examples of behaviors are included in the description of the PDC Observation System which appears in Appendix E.

Spring observation training procedures. A primary objective of the March training session was to adequately train testers so that reasonable coding reliability could be established during training and again on site before spring observation data were collected in the field. To achieve this objective, it was decided that only testers who had received training and collected observation data in the fall would be involved in spring observation training sessions.[1] This allowed spring training practices to build upon previous training and testers' observation experience in the classroom.

As in the fall training sessions, the revised observation system was introduced in a large group session. At this time, changes and revisions were highlighted and examples of the

---

[1] This decision excluded California and Michigan from the collection of observation data.

categories were provided. Small-group sessions were employed to clarify and give examples of the categories. Throughout these sessions, testers described and role-played examples of behavior they had observed in previous classroom observations. Trainers then indicated how the behaviors should be coded on the record sheet. Additional small-group sessions were used for viewing videotapes of preschool-aged children in school settings. After observers coded a two-minute segment of the tape, trainers provided feedback on how the behaviors should have been coded. Testers were asked to examine their sheets in order to detect errors. Common errors were discussed by the group; additional clarification and examples were provided for ambiguous or frequently confused categories.

As the training week progressed, testers felt more confident about their observations, and videotapes were coded with greater accuracy and consistency. At the end of the training week, reliability data were collected to assess the accuracy of testers' observations and coding.

Observation procedures. In an attempt to insure that observation data would be collected in a consistent manner across sites, guidelines and procedures for completing observations were specified in the training session and closely adhered to by observers in the field.

Before they began observations at the site, testers met with classroom teachers to explain the observation instrument and answer questions. To control for observation bias, testers completed all observations prior to administering any tests and observed only the children that appeared on their rosters. Beginning with the first three children on their roster, the testers observed each child for five-minute intervals at four different time points. Each five-minute interval was divided into 15 20-second units. These units were further divided into five seconds for observing and 15 seconds for recording. The observing and recording intervals were signaled by a portable cassette tape recorder that emitted an electronic "beep" into an earphone worn by the observer.

8

Observation guidelines emphasized that observers should alternate five-minute observations among the children. Therefore, in the classroom the observers viewed each of three children for five minutes, then repeated this observation pattern three times. Between 60 and 90 minutes were needed to complete the four observations of three children. This procedure was followed until all groups of children were observed.

The number of five-minute observations coded per child and the grouping of children varied with the number of children in the classroom, the class schedule for that particular day, and extraneous factors such as absence or special instruction outside the classroom. Testers were advised to observe during all periods of the day except outdoor play and toileting. If, however, regular classroom activities such as storytime, art or snack time were conducted outdoors, testers were instructed to observe during those times. Testers were also instructed in how to deal with situations that interrupt observations. A more detailed account of the observation procedures can be found in Interim Report III, Part A, March, 1976.

Collection of reliability data. The reliability estimates included in this report were gathered from the March training session and from on-site observations completed by tester trainers and testers at each of the sites collecting observation data. A training week reliability assessment was needed to determine how adequately testers were prepared to begin their observations in the field. To assess this, small groups of testers simultaneously watched and independently coded a thirty-minute videotape. The videotape included several examples of the behavioral categories contained in the observation system. While this assessment of reliability allowed trainers to give testers individual feedback on their coding accuracy, it was not a sufficient index of data quality since it did not replicate classroom settings. Therefore, a second set of reliability data was collected when tester trainers visited the sites for test monitoring purposes.

In collecting on-site reliability data, trainers and testers completed at least 20 minutes of joint observations on three children who were not members of the PDC or comparison samples. The trainers and testers simultaneously observed and independently coded their behavior.

Analysis of reliability data. For both sets of reliability data, the tester's coding was compared to a single trainer's coding of the same behavior events. Although the measure thus produced is not identical to a measure of inter-observer agreement, it does assess the accuracy of testers' coding as compared to a single standard criterion. This provides a basis for detecting those categories that were commonly coded unreliably by a majority of the testers. Further, analytical inferences that include these categories could take this factor into account.

Two methods for assessing coding reliability were employed. The first method computed a pairwise tester and criterion agreement estimate within categories for each 5-second observation interval. A proportion of agreement was determined using Cartwright's alpha.[1] This procedure consists of comparing, unit by unit, the codes selected by the tester with the criterion codes. Estimates were obtained for the number of times tester and criterion codes agreed and disagreed for each observation unit. The reliability figure was then computed by taking the number of times codes agreed, and dividing this number by the number of agreements plus disagreements.

A second method of reliability assessment was also used. This method examined how well testers' codes matched the criterion for the total length of a given observation period. Although the testers' codes may not agree with the criterion unit by unit, it is important that testers, after viewing a child for a specified interval, at least agree on the relative number of tallies assigned to each subcategory. To obtain this estimate, the total number of tallies assigned to a given category by one observer was divided by the total number of tallies assigned to that category by the other observer. Different inferences can be drawn from these two reliability estimates: the first, based upon Cartwright's alpha, indicates the reliability of a single observation within a category; the second, overall proportion of agreement, indicates the reliability of the total frequency of observations for a category.

Reliability results. The mean proportions of agreement and alphas within each subcategory for the training week and on-site reliability assessments are presented in Tables 15 and 16. Also included in these tables is an indication of whether observers overestimated or underestimated the frequency of a specific subcategory.

An examination of these agreement figures reveals that, in general, coding errors occurred most often in categories describing the child's behavior during child-peer and child-adult interactions. Overall proportions of agreement in these categories ranged from .00 to .91, with a mean of .71. Low reliability figures in some subcategories were, to some extent, expected since the majority of these subcategories consisted of descriptor items that had been redefined or newly added to

---

[1] Cartwright, D.S. A rapid non-parametric estimate of multi-judge reliability. Psychometrika, 1956, 21, 17-29.

Table 15

Training Week Coding Reliability

Spring 1976

Observation Cateogy		Overall Proportion of agreement	Cartwright's alpha	Direction of Error[1]
Involvement	NONINVOLVED	.74	.54	+
	SOCIAL	1.00	.93	
	NONSOCIAL	.84	.79	+
Verbal Behavior	VERBAL ENGLISH	.94	.92	+
	VERBAL SPANISH	X[2]	X	
	VERBAL COMBINED	X	X	
	NONVERBAL	.99	.88	-
Peer Interactions	NEGATIVE CONTROLS	.00	.00	-
	POSITIVE CONTROLS	.81	.50	+
	NEGATIVE ASSERTS	.43	.30	+
	POSITIVE ASSERTS	.59	.36	+
	NEGATIVE OTHER	.00	.00	-
	POSITIVE OTHER	.98	.78	+
	REQUESTS INFORMATION	.77	.53	+
	GIVES INFORMATION	.73	.64	-
	REQUESTS ASSISTANCE	.62	.58	+
	GIVES ASSISTANCE	.00	.00	-
	REQUESTS SUPPORT	X	X	
	GIVES SUPPORT	X	X	
	OTHER PURPOSES	.89		+
Adult Interactions	NEGATIVE CONTROLS	X	X	
	POSITIVE CONTROLS	.86	.69	+
	NEGATIVE ASSERTS	X	X	
	POSITIVE ASSERTS	.45	.38	-
	NEGATIVE OTHER	.00	.00	-
	POSITIVE OTHER	.85	.82	+
	REQUESTS INFORMATION	.92	.73	+
	GIVES INFORMATION	.78	.77	+
	REQUESTS ASSISTANCE	.87	.61	+
	GIVES ASSISTANCE	.11	.07	-
	REQUESTS SUPPORT	.00	.00	-
	GIVES SUPPORT	X	X	
	OTHER PURPOSES	.96	.52	-

[1] Overestimations & underestimations of individual observation categories are represented by the symbols + and -, respectively.

[2] "X" indicates that examples of this category did not occur during reliability coding.

Table 16

On-Site Coding Reliability
Spring 1976

Observation Cateogy		Overall Proportion of agreement	Cartwright's alpha	Direction of Error[1]
Involvement	NONINVOLVED	.74	.58	-
	SOCIAL	.97	.87	+
	NONSOCIAL	.98	.84	+
Verbal Behavior	VERBAL ENGLISH	.90	.71	+
	VERBAL SPANISH	.82	.80	+
	VERBAL COMBINED	X[2]	X	
	NONVERBAL	.99	.90	+
Peer Interactions	NEGATIVE CONTROLS	.50	.67	+
	POSITIVE CONTROLS	.91	.57	-
	NEGATIVE ASSERTS	.67	.80	-
	POSITIVE ASSERTS	.69	.15	+
	NEGATIVE OTHER	.67	.80	+
	POSITIVE OTHER	.90	.67	+
	REQUESTS INFORMATION	.75	.29	+
	GIVES INFORMATION	.64	.42	-
	REQUESTS ASSISTANCE	.67	.20	+
	GIVES ASSISTANCE	.55	.39	+
	REQUESTS SUPPORT	X	X	
	GIVES SUPPORT	X	X	
	OTHER PURPOSES	.90	.64	
Adult Interactions	NEGATIVE CONTROLS	--	--	
	POSITIVE CONTROLS	.88	.64	-
	NEGATIVE ASSERTS	X	X	
	POSITIVE ASSERTS	.40	.29	+
	NEGATIVE OTHER	X	X	
	POSITIVE OTHER	.97	.80	-
	REQUESTS INFORMATION	.83	.64	+
	GIVES INFORMATION	.92	.62	+
	REQUESTS ASSISTANCE	.70	.24	+
	GIVES ASSISTANCE	.88	.53	+
	REQUESTS SUPPORT	.50	.67	-
	GIVES SUPPORT	X	X	
	OTHER PURPOSES	.98	.75	-
Activity Levels	MAXIMUM	.99	.96	-
	MODERATE	.85	.90	+
	MINIMUM	.81	.85	-

[1] Overestimations and underestimations of individual categories are represented by the symbols + and -, respectively.

[2] "X" indicates that examples of this category did not occur during reliability coding.

the revised observation instrument. In addition, these categories required testers to make fine distinctions regarding the nature and purpose of the child's social interactions.

In contrast, subcategories that were a part of the observation instrument used in the fall data collection contained higher proportions of agreement. It appears that observers could accurately distinguish and code children's involvement in the classroom (Noninvolved, Social, and Nonsocial--see Appendix E for expanded definitions of terms used here). The overall proportion of agreement for these subcategories ranged from .74 to 1.00, with a mean of .88. Also, whenever a child was involved in an activity, testers could accurately code the child's verbal behavior (Verbal English, Verbal Spanish, Verbal Combined, and Nonverbal). Agreement estimates in these subcategories ranged from .82 to .99, with a mean of .93.

The fall observation data analysis found children's behavior to be highly related to their opportunities for social interaction in the classroom. In view of this relationship, it was decided for the spring analysis that children's relative frequency scores should reflect the relative amount of time they were observed in a given activity level, and findings would be reported according to the activity level of the classroom. Thus, it was particularly important that there be high proportions of agreement for items describing the activity level of the classroom (Maximal, Moderate, and Minimal). The analysis of on-site reliability data indicated that testers could accurately assess and code the opportunity for social interactions in the observed classrooms. The reliability estimates for this category ranged from .81 to .95, with a mean of .88.

Reliability of individual social interaction items. As discussed earlier, many of the subcategories included in the Peer and Adult interaction categories require the tester to make fine distinctions among the descriptor items. For example, if a child is observed requesting something from a peer, the tester not only codes that the child is requesting, but also distinguishes and codes whether the object of the request is information, assistance/materials, or emotional support. Disagreements in these categories, therefore, stem from two sources of coding error. Testers may agree at a more general level that the child is requesting something, but at a more specific level they may disagree about the object of the child's request. It was not surprising then, that the reliability figures in these subcategories were relatively lower than other subcategories.

To gain information concerning the source of coding errors, the proportions of agreement for individual descriptor items were examined. Tables 17 and 18 present the findings of this supplemental analysis. As expected, higher proportions of agreement appeared in those items representing global descriptions of children's behavior (Negative, Positive, Requests, and Gives); whereas lower agreement occurred in those items that are more specific (Control, Assert, Information, Assistance/Materials, and Support). This lower agreement may also be attributed to the infrequency of these behaviors in the classroom and the coding reliability assessments. In view of this, information and inferences derived from the analysis of these subcategories are considered tentative.

Preparation of observation data for analysis. For all observation categories, a sum of the child's behavioral incidents across the five-minute observation intervals was computed. Each child, then, had one summary score for each item on the observation instrument. Because the results from the fall data analysis indicated that children's behavior varied according to the general activity level of the classroom (Maximal, Moderate, and Minimal), a child's score should take this variation into account. Thus, as in the fall analysis, relative frequency scores for the observation variables were computed for each activity level, weighting the absolute scores by the amount of time children were observed in this activity level. For example, if the child's score on Social Involvement was .53 this would mean that the child was socially involved for 53% of the time intervals during which he or she was observed. This figure is derived by dividing the Social Involvement summary score by the number of observation units.

Results of descriptive analysis. In order to provide a summary description of the observation data, the relative frequency for each observation variable was computed for each activity level. Frequencies were then aggregated across sites.[1] Figure 4 presents the relative amount of time children exhibited Noninvolved, Social, and Nonsocial behaviors. The results indicate that children spent high proportions of their time in activities that involved objects (40%), and in activities that involved objects and/or persons (52%). These proportions were found to vary across activity levels with moderate and minimal activity levels proving most conducive to social interactions.

---

[1] The means and standard deviations for each variable by activity level are included in Appendix E.

Table 17

Coding Reliability of Individual Subcategory Items for Training Week Assessment

Spring 1976

	Subcategory Items	Overall Proportion of Agreement	Cartwright's alpha	Direction of Error[1]
Peer Interactions	Control	.76	.68	+
	Assert	.57	.41	-
	Other	1.00	.80	+
	Negative	.72	.40	-
	Positive	.99	.86	-
	Information	.76	.68	+
	Assistance/Materials	.96	.56	-
	Other Purposes	.89	.79	-
	Give	.70	.64	+
	Requests	.62	.65	-
Adult Interactions	Control	.82	.69	+
	Assert	.46	.41	+
	Other	.87	.83	-
	Negative	X[2]	X	X
	Positive	.95	.93	+
	Information	.82	.81	-
	Assistance/Materials	.78	.58	+
	Other Purposes	.91	.55	+
	Gives	.91	.77	-
	Requests	.92	.82	-

---

[1] Overestimations and underestimations of individual categories are represented by the symbols + and -, respectively.

[2] "X" indicates that examples of this category did not occur during reliability coding.

Table 18

Coding Reliability of Individual Subcategory Items for Onsite Reliability Assessment

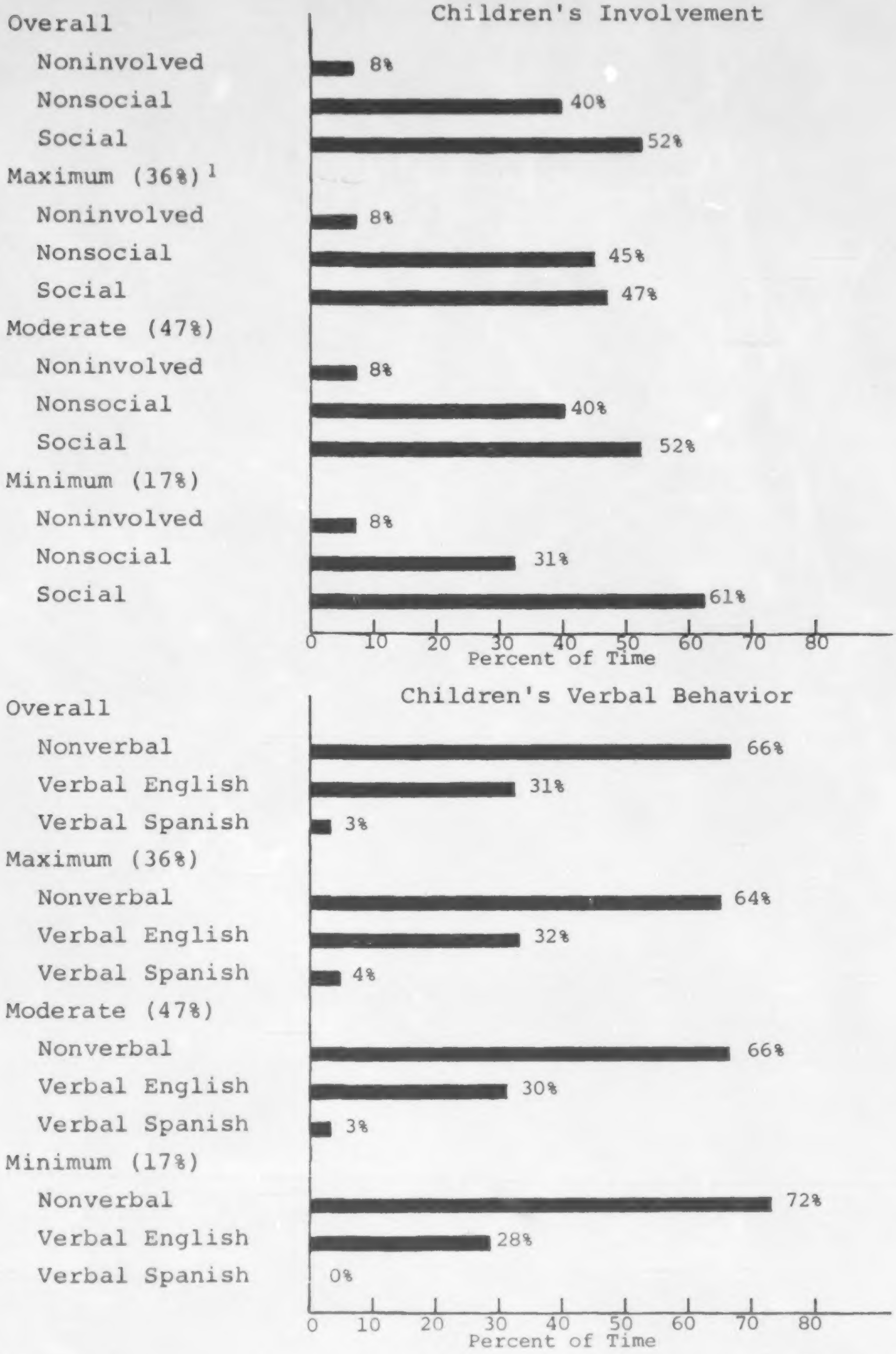
Spring 1976

	Subcategory Items	Overall Proportion of Agreement	Cartwright's alpha	Direction of Error[1]
Peer Interactions	**Control**	.88	.56	-
	Assert	.78	.25	+
	Other	.90	.68	+
	Negative	.86	.92	+
	Positive	.91	.78	+
	Information	.71	.43	-
	Assistance/Materials	.63	.38	+
	Other Purposes	.90	.64	+
	Give	.80	.46	
	Requests	.67	.27	+
Adult Interactions	**Control**	.96	.63	-
	Assert	.50	.29	+
	Other	.97	.80	-
	Negative	X	X	X
	Positive	.98	.82	-
	Information	.91	.61	+
	Assistance/Materials	.78	.44	+
	Other Purposes	.97	.74	-
	Gives	.91	.62	
	Requests	.84	.61	+

[1] Overestimations and underestimations of individual categories are represented by the symbols + and -, respectively.

[2] "X" indicates that examples of this category did not occur during reliability coding.

Figure 4
Relative Frequencies of Children's Involvement and Verbal Behavior, by Classroom Activity Level
Spring 1976
$N = 469$



[1] Percentage indicates the relative amount of time children are observed under conditions that respectively permit maximum, moderate, and mimimum opportunity for interaction.

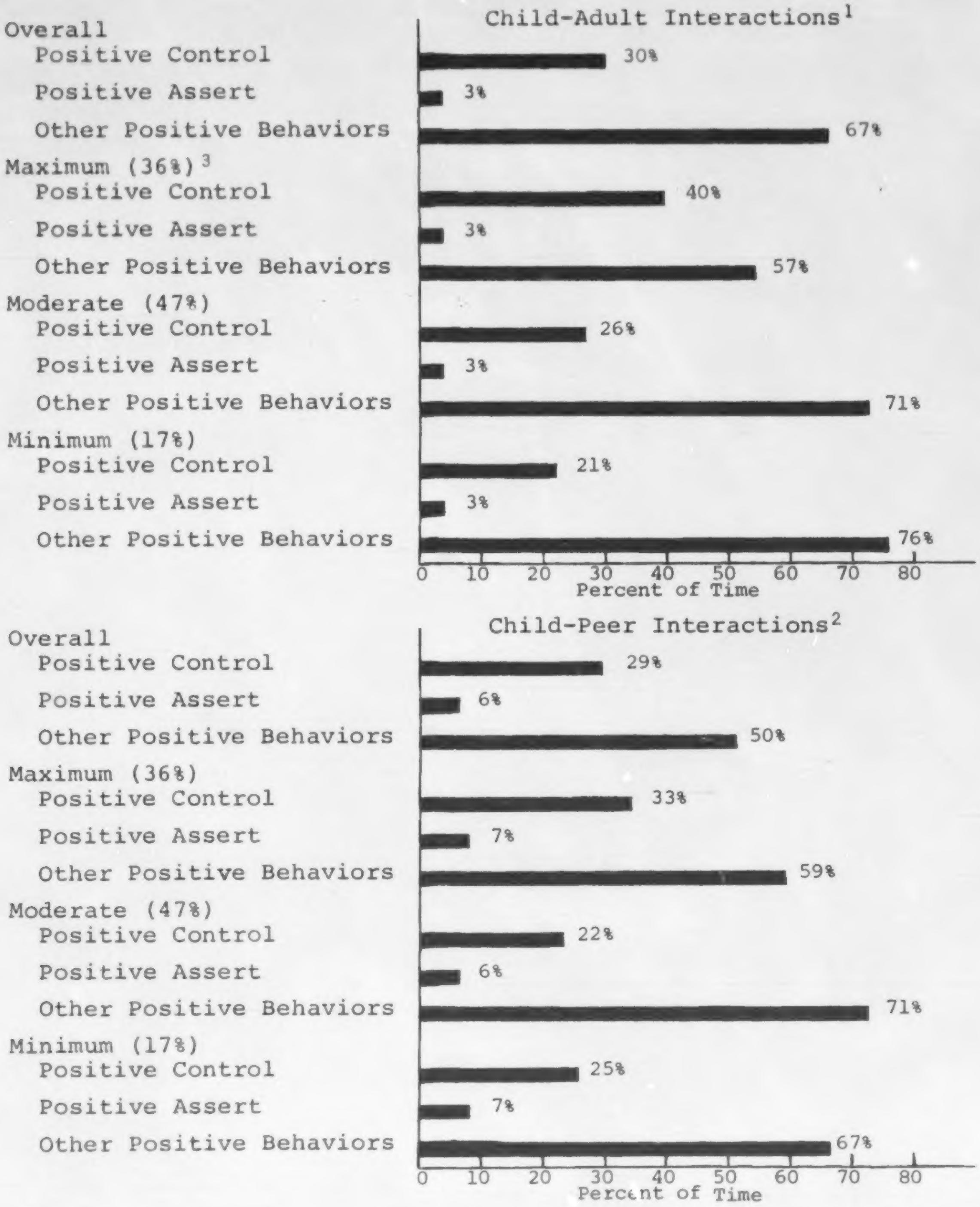
Also presented in Figure 4 are the relative frequencies of children's verbal behavior. Children were verbal less than 31% of the time, with only a slight variation across activity levels. For the Bilingual/Bicultural Demonstration site (Texas), the proportion of Spanish spoken by children ranged from 0% for minimal activity levels to 24% for maximal activity levels, with a mean of 18%.

Figures 5 and 6 display the relative frequencies of child-peer *and child*-adult interaction variables by subcategory. As in the fall, children's social interactions primarily involved peers (59%). Adult interactions occurred 32% of the time, while joint interactions with both an adult and peer occurred 32% of the time. These percentages sum to more than 110 because joint interactions resulted in multiple codings. It appears that children primarily interacted in a positive manner with peers and/or adults, for negative behaviors were exhibited less than 1% of the time. The definition of Control was slightly different than the previous fall's, and attempts to Control peers (30%) and adults (29%) has substantially increased in comparison. Spring proportions of these categories were only moderately influenced by the activity level of the classroom.

Behaviors that reflected children's attempts to resist the control of others (i.e., Assert) occurred 6% to 10% of the time relative to the activity level of the classroom. Of course, the occurrence of this behavior is related to the amount of control exhibited by others. In comparing adult and peer interactions, it appears that more Positive Asserts were emitted in the presence of adults. This suggests that adults were directing children's behavior more frequently than peers, and that children were resisting this control in a positive manner.

Finally, children exhibited high proportions of Gives behaviors (49%). An examination of the object of these interactions reveals that children provided information 45%-50% of the time, and assistance or materials 4%-5% of the time. In contrast, Requests behaviors were exhibited less than 20% of the time. These behaviors were primarily directed toward obtaining assistance or materials from peers (13%) and information from adults (15%). Patterns of children's Requests and Gives were found to fluctuate across activity levels. As activity levels changed from maximal to minimal conditions for social interactions, Gives behaviors within the context of peer interactions decreased, while Gives behaviors directed toward adults increased. High proportions of Gives Information characterized minimal activity levels. Because large group instructional activities (e.g., music, storytime) typify these

Figure 5
Relative Frequencies of Adult and Peer Interaction Variables;
Nature of Interaction by Classroom Activity Level
Spring 1976
N = 469



Child-Adult Interactions[1]

Child-Peer Interactions[2]

---

[1] Child-Adult interactions represent 32% of children's social interactions.

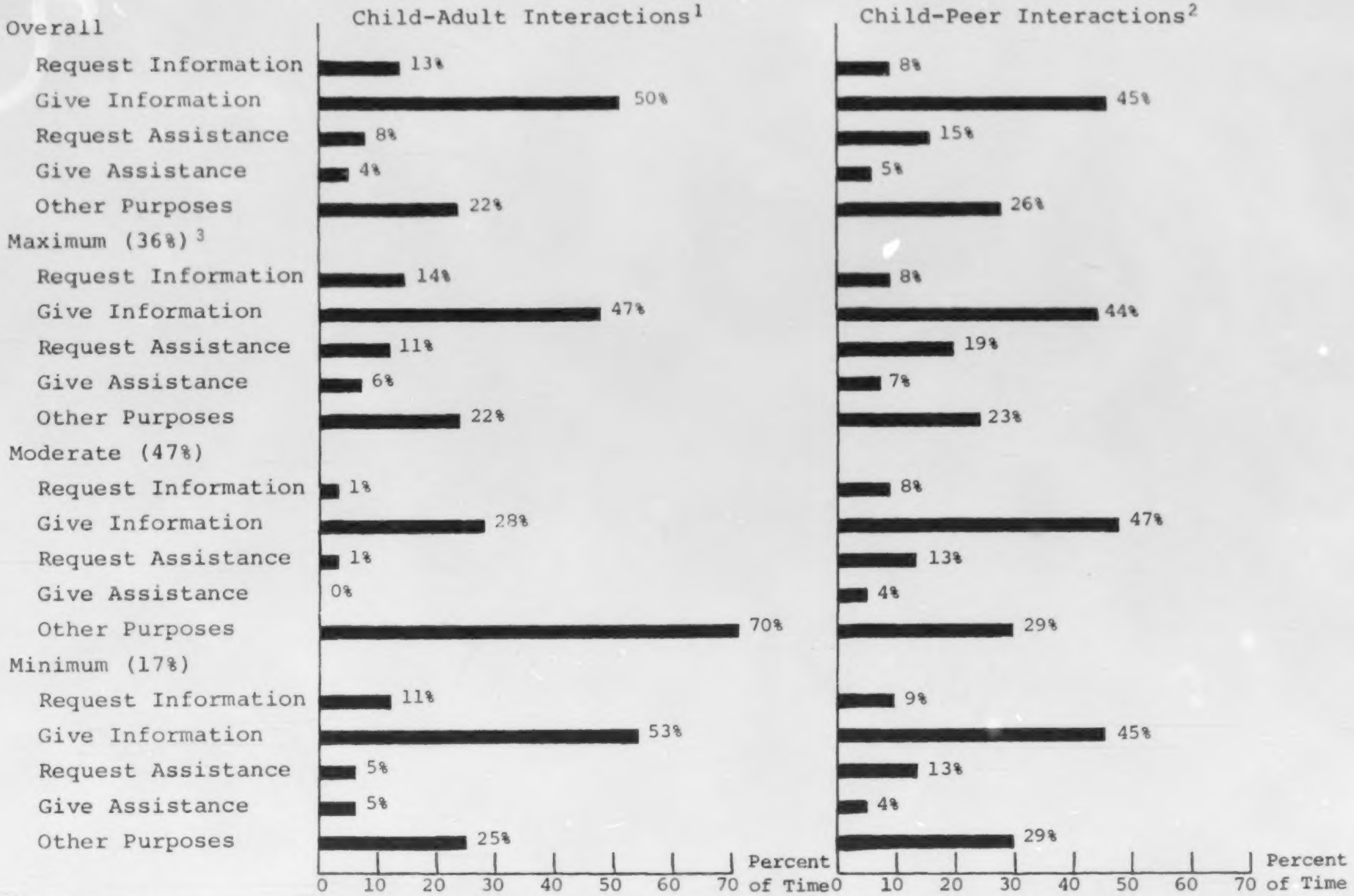
[2] Child-Peer interactions represent 59% of children's social interactions.

[3] Percentage indicates the relative amount of time children are observed under conditions that respectively permit maximum, moderate, and minimum opportunity for classroom interaction.

Figure 6

Relative Frequencies of Adult and Peer Interaction Variables; Purpose of Interaction by Classroom Activity Levels, Spring 1976

N = 469



[1] Child-Adult interactions represent 32% of children's social interactions.

[2] Child-Peer interactions represent 59% of children's social interactions.

[3] Percentage indicates the relative amount of time children are observed under conditions that respectively permit maximum, moderate, and minimum opportunity for classroom interaction.

activity levels, it is not surprising that the purpose of children's social interactions during these periods was to provide information. Moreover, this change in the pattern of children's social interactions was captured by the observation instrument, presenting supporting evidence for the validity of the instrument.

Results of correlational analysis. Based on fall findings, the correlations among observation variables were expected to show a moderate relationship between children's rate of verbal behavior and their rate of Controls, Asserts, Requests, and Gives during interactions with others. The correlations, which ranged from .02 to .15 across subcategories, suggest that children's verbal behavior was only slightly related to the strategies they used while interacting with others. In addition, child-peer and child-adult interaction categories were marked by higher correlations among the individual subcategories. Children's controlling behaviors were significantly related to the relative frequencies of Gives Information, Requests Information, and Requests Assistance. This finding suggests that controlling behaviors were generally exhibited for the purpose of directing and telling others how to do something or for getting the attention of another person to obtain information or assistance to complete a task. Thus, it appears that children's controlling behaviors were directed toward some end or goal.

It is important that the observation instrument accurately represent the social skills of children. To assess this, the observation scores were paired with relevant teacher ratings on the PDC Child Rating Scale. It was expected that a high degree of association would exist between the Child Rating Scale, which was designed to include as many as possible of the observation variables as a teacher could reasonably assess, and the actual classroom observation. An analyses of this relationship (see Table 19) indicated some agreement between selected observation items and rating scale items that assess similar dimensions of the child's behavior; but overall there appears to be low agreement between the two measures. This absence of agreement may be attributed to several factors: (1) the sample of behavior observed did not fully represent children's social skills as perceived by teachers; (2) teachers and observers may have focused on different criteria when evaluating children; and (3) teachers may have formed a generalized perception of children that affects their objectivity in rating specific behaviors.

Table 19

Correlation of Observation Variables and Corresponding Child Rating Scale Items
Spring 1976

N	Observation Variables	Child Rating Scale Items	Correlation Coefficient[1]
229	*Rate of Verbal Behavior*	*Uses words or wits to try to influence others*	.36**
230	*Rate of Verbal Behavior*	*Talks freely to children*	.18*
230	*Rate of Verbal Behavior*	*Talks freely to adults*	.17*
230	*Rate of Nonsocial Involvement*	*Works well on his/her own*	-.02
228	*Rate Positive Peer Interactions*	*Gets attention of peers appropriately*	.00
228	*Rate Positive Adult Interactions*	*Gets attention of adults appropriately*	.07
229	*Rate Positive Adult & Peer Interactions*	*Cooperates & shares with others*	-.02
229	*Rate Social Involvement*	*Shows self-confidence*	.16*
230	*Rate Social Involvement*	*Attempts to solve social problems with little adult assistance*	.09
227	*Rate of Negative Adult & Peer Interactions*	*Shows verbal dislike or hosility to others*	.09
227	*Rate of Negative Adult & Peer Interactions*	*Shows physical dislike or hostility to others*	.06
228	*Rate of Assertive Adult & Peer Interactions*	*Asserts his/her right to fair treatment*	.01
178	*Rate of Requesting Assistance from Adults & Peers*	*Attempts to solve social problems with little adult assistance*	.18*
178	*Rate of Requesting Assistance from Adults & Peers*	*Asks for needed help to do a task*	-.14
178	*Rate of Requesting Assistance from Adults & Peers*	*Asks for unnecessary help to do a task*	-.13
229	*Rate of Controlling Adults & Peers*	*Succeeds in influencing or controlling others*	.00
227	*Rate of Controlling Adults & Peers*	*Uses words or wits to try to influence others*	.12

*p<.01
**p<.001

Of additional interest is the relationship between observation items and other child measures. In examining this, correlations between the observation variables and other measures that were found to be reliable were computed. These correlations were computed for both individual and classroom means on child tests and observation variables. The correlations based on individual scores are shown in Table 20; those based on classroom means appear in Table 21. In general, the correlations based on class means are higher than those based on individual scores. Of course, with the reduced N, classroom-level analyses result in fewer significant correlations.

In contrast to the fall correlational findings, a number of observation variables showed a high degree of association with the other child measures. Higher correlations marked those categories that occurred more frequently (e.g., Social, Nonsocial, Nonverbal, etc.). For example, the amount of time that children spent in interactions with adults was significantly related to their verbal competencies as assessed by the verbal subtests. Although behavior in the individual subcategories describing children's peer and adult interactions occurred infrequently, some of these categories were moderately, and in some cases highly, related to the other child measures. In particular, children's rates of Gives and Requests were highly associated with their scores on the five Child Rating Scale factors. It is of some concern, however, that low correlations did occur within some observation categories. This may be attributable to the low reliability of the measure or the infrequency of the behavior. Alternatively, these variables may be assessing aspects of children's social competence not measured by the other instruments.

Table 20

Correlation of Observation Variables and Other Reliable Measures Based on Individual Scores[1]
Spring, 1976

Child Measure	Children's Involvement			Children's Verbal Behavior				Descriptors of Children's Peer Interactions							Descriptors of Children's Adult Interactions						Purpose of Peer Interactions				Purpose of Adult Interactions			
	rate noninvolvement	rate social involvement	rate nonsocial involvement	rate English	rate Spanish	rate combined	rate nonverbal	rate peer interactions	rate positive control	rate negative control	rate positive assert	rate negative assert	rate positive other	rate negative other	rate adult interactions	rate positive control	rate negative control	rate positive assert	rate positive other	rate negative other	rate information requests	rate information gives	rate assistance requests	rate assistance gives	rate information requests	rate information gives	rate assistance requests	rate assistance gives
BSM - English (n=349)	-.13	-.02	.09	.14	-.10	-.04	-.12	.07	.06	-.04	.02	.02	.07	-.05	.23	-.13	.05	-.02	.04	.05	.04	.04	-.10	-.05	-.08	.00	-.16	-.09
BSM - Spanish (n=62)	.07	-.20	.18	-.15	.09	-.06	.03	-.02	.04	.10	.00	.03	.15		.19	-.18		-.11	.23		-.01	-.03	-.06	.20	*-.24*	.10	*-.27*	.07
Arm Coordination (n=408)	.07	-.08	.05	*-.11*	*.12*	-.01	.06	.05	-.01	-.02	-.06	.03	-.01	-.01	-.09	.02	.09	*-.09*	.05	-.04	.06	.02	*-.09*	.03	.02	.08	-.03	.07
Arm-Leg Coordination (n=381)	-.01	*.09*	*-.09*	-.01	*.15*	-.01	-.08	*.08*	.00	-.01	.02	.01	.00	.01	-.05	-.05	*.11*	-.08	-.03	-.01	.00	-.03	-.04	.02	.03	-.04	.01	.10
Draw-A-Child (n=408)	*-.12*	*-.07*	*.13*	-.02	-.06	-.04	.06	-.04	.02	-.04	-.08	-.07	.02	-.07	*.08*	*-.03*	.04	-.04	.05	-.02	.02	.01	-.04	-.06	-.02	.00	.02	*-.11*
Verbal Fluency (n=408)	-.05	*-.11*	*.14*	.06	*-.14*	-.02	.01	-.01	.03	-.03	-.06	.03	-.06	-.02	*.12*	.12	-.02	-.07	*.14*	.01	-.02	.03	.02	-.07	.08	-.01	-.09	-.02
Verbal Memory-1 (n=339)	-.03	-.06	.08	-.04	.00	.02	.04	-.01	.07	-.02	-.05	.02	-.04	.00	.02	-.02	.04	.05	.00	.04	-.02	.01	.04	-.03	.00	.08	*-.18*	-.02
Verbal Memory-2 (n=338)	-.03	*-.08*	*.11*	-.02	.00	-.01	.02	.01	-.02	-.05	*-.11*	-.06	-.04	-.03	*.10*	*-.13*	.01	-.08	*.15*	-.05	.09	.06	-.03	-.03	.01	-.10	-.02	.11
PIPS (n=403)	.06	*-.09*	.06	-.03	*-.10*	*-.12*	*.09*	-.02	.00	-.05	-.05	-.04	.04	-.02	.06	-.02	-.03	*-.09*	.05	-.01	.00	.00	.07	*-.15*	.02	.02	-.03	-.01
CRS-1 (n=218)	-.09	*.21*	*-.19*	*.27*	-.17		*-.25*	*-.12*	-.02	-.01	-.07	.01	*.14*	-.03	.05	-.07		-.02	.07	*.15*	-.08	.08	.09	*-.17*	.05	.08	.06	-.09
CRS-2 (n=225)	-.05	*.22*	*-.22*	*.27*	.04		*-.28*	.11	.01	.01	.06	.06	.01	-.01	.02	.01		.08	-.03	-.06	*.15*	-.02	-.11	-.04	.11	-.01	.09	.00
CRS-3 (n=226)	.00	-.05	.05	.01	*-.13*		.01	.01	.00	.11	-.07	.02	*.15*	-.07	-.04	-.07		-.07	.08	.10	*-.12*	*.13*	.08	*-.11*	-.06	.09	-.02	*-.16*
CRS-4 (n=229)	-.03	*.12*	*-.12*	.09	-.07		-.08	-.06	-.10	.03	-.09	.10	*.21*	-.06	-.04	-.02		-.08	.03	*.20*	-.08	*.15*	*.14*	*-.17*	-.08	*.16*	*.14*	*-.14*
CRS-5 (n=229)	*-.13*	.08	-.02	.10	-.04		-.10	-.02	.03	-.04	-.07	-.09	.07	-.03	.10	-.03		-.04	-.02	*.20*	-.10	.03	.09	-.08	.01	.00	.04	-.06
POCL-1 (n=469)	*-.08*	.02	.02	-.02	*.18*	.02	*-.08*	-.03	.02	-.03	.00	.00	-.02	-.04	.04	-.01	.01	-.08	.04	-.02	-.05	.08	.05	-.05	.04	-.06	.02	-.08
POCL-2 (n=469)	*-.12*	*.10*	-.04	*.13*	.07	.02	*-.19*	.06	-.02	-.05	.01	-.06	.01	-.01	.13	-.05	-.01	.01	.05	.03	-.06	.09	-.03	-.07	.06	-.03	.01	.02

---

[1] Significant correlations at the .05 level and above are presented in italics.

Table 21

Correlation of Observation Variables and Other Reliable Measures Based on Classroom Means[1]
Spring 1976

Child Measure	Children's Involvement			Children's Verbal Behavior				Descriptors of Children's Peer Interactions							Descriptors of Children's Adult Interactions						Purpose of Peer Interactions				Purpose of Adult Interactions			
	rate noninvolvement	rate social involvement	rate nonsocial involvement	rate English	rate Spanish	rate combined	rate nonverbal	rate peer interactions	rate positive control	rate negative control	rate positive assert	rate negative assert	rate positive other	rate negative other	rate adult interactions	rate positive control	rate negative control	rate positive assert	rate positive other	rate negative other	rate information requests	rate information gives	rate assistance requests	rate assistance gives	rate information requests	rate information gives	rate assistance requests	rate assistance gives
BSM - English (n=38)	-.06	-.20	.25	-.11	.20	.19	-.04	.03	.10	-.14	.10	.11	.25	-.23	.38	-.13	.03	.31	.06	.03	.28	-.18	-.39	.24	-.29	.10	-.33	-.04
BSM - Spanish (n=7)	.06	-.48	.49	-.39	.20	-.40	.13	.46	.35	.13	.20	.25	-.56		.06	-.07		*-.62*	.35	-.09	-.40	-.67	.27	.39	-.75	.10	.15	.20
Arm Coordination (n=41)	.10	-.21	.18	-.25	.07	.04	.26	.04	-.21	.01	*-.30*	.02	*-.35*	-.09	-.23	-.22	-.06	-.28	.27	.00	.18	.02	-.21	.09	.01	*.39*	*-.33*	.01
Arm-Leg Coordination (n=41)	-.24	.11	.00	-.19	*.33*	.14	-.04	.07	.01	*.33*	-.07	.23	-.16	-.07	-.13	.25	-.04	-.13	-.21	-.06	-.26	-.02	.04	.07	.26	-.22	-.11	*.39*
Draw-A-Child (n=46)	.12	*-.42*	*.42*	*-.35*	-.07	-.12	*.51*	-.12	.00	-.21	-.26	-.18	-.20	-.22	.10	*-.41*	-.01	-.18	*.43*	.10	.26	-.03	*-.30*	-.03	-.14	.24	-.12	*-.30*
Verbal Fluency (n=46)	.05	*-.30*	*.30*	-.16	*-.29*	-.25	*.45*	-.15	-.18	*-.30*	*-.35*	-.14	-.28	-.12	-.07	*-.30*	-.10	*-.40*	*.37*	-.10	.15	-.02	-.10	-.09	-.01	.27	-.15	*-.40*
Verbal Memory-1 (n=339)	.06	-.21	.21	-.23	.17	.13	.16	.08	-.07	.02	.05	.01	*-.41*	-.04	-.02	-.17	-.01	.07	.14		.10	-.05	*-.48*	.20	-.17	.27	*-.38*	.14
Verbal Memory-2 (n=338)	-.06	-.26	.33	*-.43*	.08	.07	*.49*	-.24	-.19	-.24	*-.33*	-.18	-.21	-.25	.00	-.24	-.05	-.14	.26	-.03	.17	.11	-.22	-.14	.04	.21	*-.31*	-.41
PIPS (n=46)	.22	*-.29*	.21	-.20	-.24	*-.35*	*.46*	-.13	-.15	-.03	*-.31*	-.03	-.18	-.12	-.13	-.22	-.02	*-.32*	*.28*	.03	.21	-.08	.04	*-.30*	.07	.14	-.03	-.36
CRS-1 (n=28)	.04	.06	-.09	-.10	-.19		.17	-.23	-.01	-.19	-.26	-.25	*-.54*	-.11	-.02	.30		.00	*-.30*	.10	-.03	.12	*.40*	*-.36*	.03	.14	.00	-.13
CRS-2 (n=29)	-.04	.14	-.13	-.08	.17		.04	-.13	.11	-.16	-.03	.02	-.11	.07	.04	.33		-.07	*-.32*	.11	.23	-.07	-.05	.06	.15	*-.43*	.28	.19
CRS-3 (n=29)	-.02	-.08	-.10	.24	-.30		-.17	.12	-.08	-.04	.08	.18	.25	-.09	.05	-.25		.11	.24	.04	-.12	.00	.04	-.04	-.07	*.40*	*-.40*	-.09
CRS-4 (n=29)	.14	.02	-.10	-.09	-.15		.15	-.07	-.15	-.20	-.08	.26	*.56*	-.19	-.16	*.31*		.04	*-.38*	.26	-.04	*.42*	*.37*	*-.33*	-.18	.29	.31	-.14
CRS-5 (n=29)	.01	-.11	.11	.00	-.24		.07	-.15	.12	-.17	-.04	*-.38*	*.48*	-.11	.29	.13		.26	-.17	.21	-.08	-.08	.12	-.10	-.11	-.05	-.05	.07
POCL-1 (n=47)	-.18	.01	.08	-.20	*.32*	.11	-.01	.00	-.04	-.08	-.09	.14	.18	-.20	-.02	-.06	.03	-.23	.10	.00	-.23	.33	.09	-.10	-.07	.13	.19	-.07
POCL-2 (n=47)	-.22	.10	.00	.03	.10	.06	-.11	.13	.03	-.11	.17	-.03	.17	-.10	.01	-.15	-.03	-.11	.17	.05	-.09	.22	.00	-.20	-.02	.02	.08	-.09

---

[1] Significant correlations at the .05 level and above are presented in italics.

Site Characteristics

Comparability of PDC and Comparison Groups in Michigan

Due to communication problems between the testers and the field coordinator, the fall 1975 data collection in Michigan was incompleted, and the comparability analysis performed on data from other sites could not be completed for Michigan. Using the spring 1976 data, the PDC and comparison groups were compared both on background characteristics and on test scores. The results are shown in Table 23. Certain significant differences appeared between the two groups in terms of their background characteristics. The PDC group had a smaller proportion of black children (66% vs. 98%), had more siblings per child (2.7 vs. 1.9, a difference equal to .57 standard deviation), their mothers were more likely to be employed, and they were more likely to have fathers present at home. The two groups also differed significantly on 2 of 11 tests. On the Bilingual Syntax Measure-English, the PDC children scored slightly less than a point higher (.57 standard deviation); and on the Preschoool Interpersonal Problem-Solving Task the PDC children scored slightly more than a point higher (.65 standard deviation).

In absolute terms and in respect to other sites, the comparability of groups in Michigan could be improved. None of the group contrasts performed on data gathered in the fall from other sites revealed as many differences between PDC and comparison children (see Interim Report III, Part A, p. 53). However, Michigan, like other sites, is attempting to select better-balanced groups of children for enrollment in fall 1976 Head Start classes, and this may alleviate the problem.

Recommended Modification of the Evaluation Design for Georgia

In Interim Report II, Part C (June 30, 1975) a cross-sectional comparison group design was proposed for the Georgia site. A cross-sectional design is an alternative to a contemporaneous comparison group design: in the case of PDC Head Start children in Georgia, the comparison group consists not of other Head Start children, but of children who are now in grades K, 1, 2, and 3. Samples of children from each of these levels, Head Start through grade 3, were tested in the fall of 1975 and the spring of 1976; in ensuing years, as the PDC children progress from Head Start through the higher grades, they will be retested each fall and spring. It was originally proposed that their

Table 22

Michigan: Magnitude of Differences for Variables on which Groups Were Found Unequal

		PDC	Comparison	Difference*	Probability
Ethnicity	N	32	50		
	% Black	66	98	-32	
	% Hispanic	03	0	03	
	% White	31	2	29	.0003
Number of Siblings	N	32	49		
	Mean	2.72	1.87	.57z	
	SD	1.80	1.27		.02
Mother Employed?	N	14	25		
	% Yes	21	04	17	
	% No	79	96	-17	.08
Father at home?	N	31	18		
	% Yes	32	11	21	
	% No	68	89	-21	.10
BSM-English	N	31	49		
	Mean	4.16	3.29	.57z	
	SD	1.72	1.40		.02
Preschool Interpersonal Problem-Solving Task (PIPS)	N	29	49		
	Mean	3.38	2.31	.65z	
	SD	1.74	1.60		.007

*"Difference" is equal to the PDC figure minus the comparison figure. In the case of metric variables, means are reported as raw scores while the difference is reported as a z score.

scores be compared with the scores obtained by children from the appropriate grade level of the cross-section. That is, when the PDC children reached kindergarten, their current test scores were to have been compared with the scores of children who were in kindergarten in 1975-76, the following year their scores were to have been compared with those of children who were in grade 1 in 1975-76, and so forth through the successive grades. The main reason for proposing such a design for Georgia was that no suitable contemporaneous comparison group could be located at that site. It was also felt that a tryout of this design would be of interest to planners of future Head Start evaluations. While a cross-sectional design still seems feasible and valuable, recent findings of PDC-comparison group differences in Georgia require that certain alterations be made in the procedure originally conceived.

PDC-comparison group imbalances reflected by fall 1975 data. Analysis of the fall 1975 data revealed that, as in a number of other sites, the PDC and comparison children in Georgia are dissimilar in several respects: the PDC group has a greater proportion of black children and apparently has a greater proportion of children with prior preschool experience. It is likely that the reason for these differences is that Head Start in the Georgia site serves a population that is not identical to the population served by the elementary school. Although efforts are being made by the local staff to recruit a Head Start group for 1976-77 that more closely resembles the elementary school sample, it is unlikely that the samples can be matched completely if the populations are as dissimilar as they seem.

Constitution of baseline scores by means of regression analysis. If the PDC and comparison children were compared directly in later years for purposes of program evaluation, it would be difficult to determine just what portion of any observed group difference was due to the program itself, since the groups can be presumed to have differed originally. Thus it is necessary to take the original differences into account in order to assess program effects accurately. According to present plan, this will be accomplished by means of regression analysis. Based upon knowledge (gained in this year's fall and spring testing) of the relationship between scores on the various tests in the PDC battery and predictor variables such as age, prior school experience, ethnicity, sex, and family size, scores at later time points can be predicted for each child in the PDC group entering Head Start in the fall of 1976. Evaluation of the effects of participation in PDC will then be accomplished by comparison of predicted and actual scores[1]: to the extent that PDC succeeds in promoting development, children's actual test scores in later years should exceed their predicted scores--the predictions being based upon the measured status of children who did not participate in PDC (the K-to-3 cross-section).

Descriptive Characteristics of Centers and Schools

In fall 1975, each of the sites provided a summary of the number of classes, teachers, aides, and students to be found at each level from Head Start through grade 3 within PDC and comparison centers and schools. These data were not tabulated in time for the last report, and so are presented here in Appendix F, Center and School Characteristics. No analysis of these data is contemplated at present, but they may prove useful in the future for descriptive or explanatory purposes.

---

[1] See Data Analysis Procedures, Step 5, in Chapter II for an illustration of the use of this technique.

Analysis of Attrition Trends

In spring 1975, during Year I of PDC, all sites were asked by High/Scope to submit attrition figures for local schools that could be considered similar to the designated PDC and comparison schools. These estimates provided the basis for calculation of the sample sizes required for purposes of evaluation at each site. One logical fault in the attrition estimates, though, was that they were based upon trends within the elementary school population, which may arguably differ from trends within the population of children who attend Head Start. Thus in the early spring of 1976, sites were asked to identify the children who had been enrolled since 1971-72 in what are now PDC and comparison Head Start centers, then to determine how many of the original children remain in PDC and comparison schools respectively. This information is more relevant, at least in theory, to the question of how many of the children presently enrolled in Head Start can be expected to remain until the end of grade 3. The approach met with several difficulties, however: (1) many of the centers now operating didn't exist in 1971-72--some didn't open until 1974-75--so no long-term data were available; (2) for a number of the centers that did exist, enrollment records were unavailable or were confusingly incomplete (for example, it was sometimes impossible to tell from a roster whether a given child was two, three, four, or five years old at the time of enrollment, thus it was difficult to know whose class roster to check for the following year to determine whether or not the child was still present; (3) even if the center existed in earlier years and accurate rosters could be found, Head Start-to-elementary school feeding patterns had changed in many locales, raising the problem of deciding just which schools a child had to be in to be considered "present," i.e., not a victim of attrition. These problems were generally overcome or taken into account to the extent that it was possible to develop a second estimate of attrition for almost all sites on the basis of the new data collected.

From the data gathered in 1975 and 1976, two independent estimates of annual retention rate were computed by averaging across the years for which data were available. (Annual retention rate is the proportion of the preceding year's group remaining in any given year--this proportion tends to remain fairly constant across grades.) Once the annual rate has been estimated, it is possible to compute retention over four years (Head Start to K, K to 1, 1 to 2, and 2 to 3) by raising the coefficient to the fourth power.

The figures obtained for each of the eleven sites remaining in the evaluation are shown in Table 23. The first two columns of the table show the estimates of annual retention rate. The third column shows the actual proportion of children in the testing samples who remained in school from the time of first testing in the fall of 1975 to the time of second testing in the spring of 1976 (an interval of about six months). Columns four and five show projections of the total percentage of children from an original PDC or comparison Head Start group likely to still remain in a respective PDC or comparison school through grade 3. This figure, based upon a four-year projection, may over-estimate retention somewhat, since it extends only to the beginning of what is actually the child's fifth year in school. (Testing for the proposed longitudinal PDC evaluation will continue almost until the end of the fifth year.) The sixth column in Table 23 offers an empirical check on the accuracy of the four-year retention estimate: this column shows the proportion of children who have actually remained in what are now PDC or comparison schools since their entry into Head Start in 1971-72.

It is obvious that the various estimates of retention differ to lesser or greater degrees, probably due to error in the data on which they were based; yet in averaging across sites, the different figures are in fairly close agreement: the two independent estimates of annual retention are 82% and 81%, while the actual figure for a six-month period is 88%; the two projections of four-year retention are 46% and 42%, while the actual figure for the groups that entered Head Start four years ago is 42%; and the two projections of the mean number of children from Head Start testing samples likely to remain until grade 3 are quite close--28 and 24 respectively.

Interpretation of attrition data. It is likely that the 1976 data are the more accurate because they were collected according to a more precise design than was employed for collecting the 1975 data. Thus estimates based upon the 1976 data should be preferred for decision-making purposes. A second consideration to keep in mind when interpreting the data is that factors other than departure from a target school will contribute to attrition from the testing sample; the two main factors are likely to be extended absence from school at testing time (e.g., due to illness) and refusal to cooperate (on the part of the child, the child's parents, or school officials). Table 24 shows that of 851 children who were to have been tested in spring 1976, about 10% moved away before or during the testing period and 3% more could not be completely tested for reasons other than departure

Table 23
Summary of Attrition Statistics

	Annual retention rate (percent)		Percent fall-spring retention in 1975-76 testing sample	Projected percent retention over 4 years		Actual percent retention from Head Start to grade 3 for 1971-72 group	Expected 1976-77 Head Start enrollment	Projected no. of children remaining until grade 3 (1980-81) based on...	
	1975 Data	1976 Data		1975 Data	1976 Data			1975 Data	1976 Data
CALIFORNIA									
PDC	95	?	81	82	?	?	47	39	?
Comp	97	79	81	89	39	21	57	51	22
COLORADO									
PDC	86	?	90	55	?	?	68	37	?
Comp	77	?	93	35	?	?	68	24	?
CONNECTICUT									
PDC	75	96	89	32	85	59	60	19	51
Comp	81	99	98	43	96	79	60	26	58
FLORIDA									
PDC	82	92	88	45	72	100[a]	45	20	32
Comp	73	96	86	28	85	100	37	10	36
GEORGIA									
PDC	80	81	85	41	43	59	60	25	26
IOWA									
PDC	84	77	73	50	35	26	60	30	21
Comp	80	74	89	41	30	45	65	27	20
MARYLAND									
PDC	64	74	92	17	30	39	70	12	21
Comp	75	69	93	32	23	07	60	19	14
MICHIGAN									
PDC	78	67	91	37	20	35	75	28	15
Comp	92	93	86	72	75	68	60	43	45
TEXAS									
PDC	?	86	95	?	55	65[a]	45	?	25
Comp	?	86	88	?	55	65[a]	45	?	25
UTAH									
PDC	80	74	86	41	30	35	65	27	20
Comp	78	53	88	37	08	11	65	24	05
WASHINGTON									
PDC	69	65	87	23	18	21	60	14	11
Comp	68	55	89	21	09	20	100	21	09
W. VIRGINIA									
PDC	86	73	83	55	28	?	45	25	13
Comp	93	92	88	75	72	62[b]	45	34	32
MEAN ACROSS ALL SITES	82[c]	81[c]	88	46	42	42	61	28	24

Note: A question mark indicates missing or insufficient data.

[a] No 1971-72 group. Figure shown is a projection for the 1972-73 group (oldest group for which data are available).

[b] No 1971-72 group. Figure shown is a projection for the 1973-74 group (oldest group for which data are available).

[c] Mean is weighted to take account of the geometrically greater contribution of the larger numbers to retention.

Table 24

Accounting of Complete, Partially Complete, and Incomplete Testing, Spring 1976

	Number	% of Total
Children scheduled to be tested in spring, 1976:	851	100.0
Children for whom no tests at all were received due to...		
terminated school enrollment:	85	10.0
child's refusal of cooperation:	4	.5
continual absence:	2	.2
parental refusal of cooperation:	1	.1
non-receipt of tests reportedly completed:	1	.1
other:	3	.4
Subtotal:	96	11.3
Children for whom only a partial set of tests was received due to...		
terminated school enrollment:	3	.4
child's refusal of cooperation:	8	1.0
continual absence:	6	.7
non-receipt of tests reportedly completed:	2	.2
Subtotal:	19	2.3
Number for whom a complete set of tests was received:	736	86.4
Total accounted for:	851	100.0

from school. It is probable that these difficulties will continue over the term of a longitudinal study, and may even become more serious if children, parents, or school officials grow resistant to repeated testing (it is not hard to imagine some resistance--in comparison schools particularly).

With all this in mind, the proportion of PDC and comparison Head Start children who remain in PDC and comparison schools respectively through the end of grade 3 can be reasonably estimated at 35-40%, averaged across all the sites included in the present analysis. This means that with an initial group size of 61 (the mean expected fall enrollment), the final group size at the end of grade 3 is likely to be 21-24. There will certainly be site-to-site variation around this figure--perhaps the average will even be somewhat higher due to efforts on the part of project staff to recruit Head Start children who have a high probability of remaining--but the prospect of achieving an ultimate sample of 30 children per group (the original goal) seems small. Most sites are now at maximum enrollment, precluding the enlargement of their present samples. The implications of these findings are discussed in the Conclusions section of this report.

Table 25

Summary of Findings on Test Characteristics

Child Measure	Reliability	Validity	Sensitivity to change	Suitability for higher grades	Relevance to social competence	Recommendations	Comments
BSM-English	✓	✓	?		✓*	Retain	Examine sensitivity to change and suitability for Spanish speakers.
BSM-Spanish	✓	✓	?			Retain	Examine sensitivity to change and suitability for English speakers.
WPPSI Block Design	✓	✓	✓			Retain	To be used only in Head Start year as covariate measure.
Conceptual Grouping	X		✓	X		Delete	Inadequate reliability.
Say and Tell	X			X		Delete	Inadequate reliability.
Verbal Fluency	✓	✓	✓	✓	✓	Retain	Acceptable in all respects.
Verbal Memory-1	✓*	✓*	✓	X	✓*	Retain with modification	Revise items to increase score variance among older children.
Verbal Memory-3	✓	✓	✓	✓	✓	Retain	Acceptable in all respects.
Arm Coordination	✓	✓*	✓	✓	?	Retain	Reliability, validity, and relevance to social competence may be marginal.
Draw-A-Child	✓*	✓*	✓	X	✓*	Retain with modification	Revise scoring system to increase score variance among older children.
Leg Coordination	X		X	X		Delete	Inadequate reliability.
POCL	✓	✓		✓	✓	Retain	Acceptable in all respects examined.
CRS	✓			✓	✓	Retain	Acceptable in all respects examined.
Opposite Analogies	✓				X	Delete	Expense of administration not justifiable.
Do You Know..?	✓				✓	Delete	Expense of administration not justifiable.
PIPS	✓	✓	✓	✓	✓	Retain	Acceptable in all respects.

✓ = Acceptable

? = Questionable

✓* = Provisionally Acceptable

X = Not acceptable

Blank = Not analyzed

IV

CONCLUSIONS

Summary of Findings with Regard to Instruments

The primary purpose of the spring 1976 data analysis was to determine the adequacy of the instruments included in the PDC battery. Judgments of adequacy were based on considerations of:

- reliability;
- validity;
- sensitivity to change;
- suitability for older children; and
- relevance to social competence.

Table 25 summarizes the results of analyses that focused on each of these points. On the basis of these analyses, it is recommended that the instruments listed below be retained or deleted as indicated.

Retain

Bilingual Syntax Measure-English
Bilingual Syntax Measure-Spanish
WPPSI Block Design
Verbal Fluency
Verbal Memory-1 (with modification in content)
Verbal Memory-3
Arm Coordination
Draw-A-Child (with modification in scoring)
Pupil Observation Checklist
PDC Child Rating Scale
Preschool Interpersonal Problem-Solving Test (PIPS)

Delete

Conceptual Grouping
Say and Tell
Leg Coordination
Opposite Analogies
Do You Know...?

Justification of Recommendations

BSM-English. Although the suitability of this instrument for older children and for children who are not English-dominant is open to question (data were available from only a small number of the latter group, precluding confident analysis), it is acceptable according to most criteria for English-dominant Head Start children. The indications of the instrument's sensitivity to change are ambiguous, requiring further analysis.

BSM-Spanish. This version of the BSM appears suitable for Spanish-dominant Head Start children except in its sensitivity to change, which is questionable. No BSM-Spanish data are available for older children, and the number of English-dominant children who completed it is too small to permit analysis; thus conclusions cannot be confidently extended beyond the group of Spanish-dominant Head Start children.

WPPSI Block Design. This instrument was selected for inclusion in the battery to provide a measure of early cognitive aptitude that could later serve as a covariate for analysis of other data, thus it is to be administered only during the first testing session of the Head Start year. For this purpose it appears acceptable in all respects.

Conceptual Grouping. As in fall 1975, this instrument suffered from low reliability. Although comparison of fall and spring scores revealed significant change, the imprecision of scores seems too large to be tolerable.

Say and Tell. This instrument, too, failed to meet reliability criteria, and is judged unsuitable for retention.

Verbal Fluency. This instrument passed all analytic screenings without difficulty.

Verbal Memory-1. For Head Start children, scores on this scale are reliable, valid, and apparently relevant to social competence, but none of these characteristics held for older children. The repetition of short word strings evidently becomes a simple task for older children, thus the resulting scores have little discriminatory power above the Head Start level. It is proposed that the task be restructured to increase score variance among older children.

Verbal Memory-3. The task presented by this scale, story repetition, yielded scores that were acceptable in all respects for children from Head Start to grade 3.

Arm Coordination. The relationship of this measure to social competence appears tenous, and indications of its validity are inconsistent, but since it is one of the few available indices of psychomotor performance, it is recommended that it be retained for future study.

Draw-A-Child. The scoring of this instrument presently involves assignment of points for the presence (versus absence) of body parts; this procedure yields scores that differentiate among children at the Head Start level, but not among older children (who have little difficulty drawing a complete body). An alternative scoring procedure is available that attends to the quality of the drawing rather than simply to the presence of elements, and use of this procedure is likely to permit useful differentiation among older children.

Leg Coordination. The reliability of this measure is invariably low, making it unacceptable for inclusion in the battery.

Pupil Observation Checklist. This instrument was found suitable in all the characteristics examined.

PDC Child Rating Scale. This instrument, too, appears acceptable in all the characteristics examined.

Opposite Analogies. The reliability of this instrument seems acceptable, although no convincing relationship with "social competence" criteria was found. Opposite Analogies was among the instruments offered to sites for optional inclusion in local batteries, and has been administered at only two sites. Its continued use is judged unwarranted on both theoretical and economic bases.

Preschool Interpersonal Problem-Solving Test (PIPS). This instrument appears to meet all criteria for inclusion in the battery, and may be especially useful as a key indicator of "social competence".

103

Do You Know...? This instrument also was administered at only two sites, and although it is acceptable according to the criteria used for other tests, the expense of its continued administration seems difficult to justify. Thus, it is recommended that it be deleted entirely from the battery.

Additional Instrumentation Recommendations

PDC Classroom Observation System. Although the Observation System was found in the analysis reported here to possess sufficient reliability and concurrent validity to make it potentially useful as an instrument for describing the characteristics of individual children, two factors argue for its use as an instrument for describing classroom-level characteristics; these factors are (1) measurement stability and (2) cost in time and funds.

When a child is observed for four five-minute periods in a single morning or afternoon, it is possible, perhaps likely, that the behavior sampled in that span represents only a fraction of the child's general behavior pattern--the fraction that was exposed under the classroom conditions prevailing at the time of observation. Under other classroom conditions, the child might conceivably have behaved very differently. But even if observation were continuous and exhaustive, it would still be observation of behavior in the school environment, which is only one of many environments in which children exercise social competence. Thus, even if it were possible for the observer to remain in the classroom as long as necessary to obtain a representative view of an individual child's school behavior, the view might not be broad enough to permit generalization to the child's behavior outside of school. It seems more realistic, therefore, to regard the Observation System as a measure of children's behavior in the classroom, restricting generalization to that context. However, in view of the low magnitude of the relationships found between Observation System variables and other measures available for children in the testing sample, it may be that the 20-minute sample of classroom behavior from which observation variables are constructed does not provide a stable basis for individual measurement.

The cost of the Observation System is an important additional consideration: it takes up to half as long to complete the observation of a single child as it does to give the child all the other tests in the battery, and analysis of the data produced is proportionally even more expensive because of the multiple observation cycles and the number of variables involved.

However, if the purpose of the Observation System were to become that of characterizing classes rather than individual children, total observation time could be reduced by reducing the length of the period during which each individual is observed; even with this reduction, the resulting measures, being based on more thorough sampling, can be expected to be more reliable with respect to classroom "personalities" than they are now with respect to individual children. (The larger the number of instances upon which a generalization--a score, in this case--is based, the more reliable the generalization.)

Lest it be thought that this same reasoning applies to the tests included in the battery, it is important to note how they differ from the Observation System. The tests are administered under conditions that are, as far as possible, the same for all children rather than varying with time and location; the tests are designed to elicit numerous instances of the behavior to be measured for greatest measurement reliability, rather than depending upon the spontaneous and possibly infrequent occurrence of the behavior; and the tests can be administered quickly and economically while the Observation System, used as a measure of individual behavior, is lengthy and expensive.

It is proposed, therefore, that the Observation System be regarded in the future as a measure that focuses on the nature of intact classes rather than on the nature of children. The resulting measures may prove to be of considerable value in contrasting the activities of PDC and comparison classrooms and in documenting the degree to which various aspects of the PDC program are visibly implemented in PDC centers and schools. In addition to providing measures of implementation, the Observation System, used at the classroom level, may also prove valuable to a classroom-level analysis of program impact.

PDC Child Interview (Faces). Faces, a measure of attitude toward school, was administered to children in Head Start through grade 3 during the spring testing period. Based on the number of Head Start children who did not understand the instructions (in the judgment of the testers), it is recommended that Faces not be used with Head Start children.

Response distributions (see Appendix D) for the higher grades indicate that the measure may be useful by first grade. When analysis of the data is completed, further recommendations will be made.

Preschool Productive Language Assessment Task (PPLAT). A pilot preschool version of the High/Scope Productive Language Assessment Task was administered to Head Start children in two sites during the spring testing period. Because analysis of the data from this testing has not been completed, no recommendations for its general use can be made at this time. However, because of its potential utility as a measure of language development, the present plan is to administer it at the two bilingual/bicultural demonstration sites (California and Texas) in January 1977. The outcome of the analysis presently underway may dictate revision of this plan; but if the PPLAT proves to be acceptably reliable and valid, it will provide a baseline measure of language development against which later development can be judged using more venerable instruments.

Summary of Findings with Regard to Attrition

Judging by the attrition data gathered a year ago and again this spring, it is likely that no more than a few sites, and possibly none at all, will retain a sufficient number of children from the original testing samples to permit analysis of PDC's effects within individual sites at the end of Cohort 2's third-grade year. The alternatives that immediately suggest themselves are (1) perform summative site-by-site analyses at an earlier point in PDC's history (before attrition has reduced group size below the requisite number), (2) depend upon analysis of data aggregated across all sites at the time when Cohort 2 completes grade 3, or (3) combine alternatives 1 and 2, performing site-by-site analyses each year until samples become prohibitively small, then turn to analyses of aggregate samples.

It should be kept in mind, however, that even the alternative of aggregation will require substantial samples from each site in order to permit adequate statistical control of factors other than the PDC program itself that affect children's performance. Thus selection of this alternative would not warrant relaxation of the sample size requirements previously established.

APPENDIX A

Descriptions of the Measures In the Spring Battery

Descriptions of the Measures in the Spring Battery

	Order of Administration[1]
• Social-Emotional Measures	
PDC Classroom Observation System	1
PDC Faces Interview	9
Preschool Interpersonal Problem-Solving Test (PIPS)	12
Pupil Observation Checklist (POCL)	15
• Psychomotor Measures	
Arm Coordination [McCarthy Scales of Children's Ability (MSCA)]	5
Draw-A-Child (MSCA)	6
Leg Coordination (MSCA)	13
• Cognitive and Language Measures	
Bilingual Syntax Measure (BSM)	2
Block Design (WPPSI)	(10)
Conceptual Grouping (MSCA)	3
Do You Know...? (CIRCUS)	(8)
Opposite Analogies (MSCA)	(7)
Say and Tell (CIRCUS)	11
Verbal Memory (MSCA)	14
Verbal Fluency (MSCA)	4
• Other Measures	
Adult Language Check	
Attrition, Handicap and Attendance Information Sheet	
PDC Child Rating Scale	

Each of these measures is described briefly below. For a more extensive review, see Interim Report II, Part B: Recommendations for Measuring Program Impact (1975).

---

[1] As noted in the text, the battery was administered in two or sometimes three sessions. Parentheses mark measures that were not administered at all sites.

PDC Classroom Observation System (High/Scope Foundation, unpublished). The PDC observation system was developed to provide information about children's classroom behavior along dimensions pertinent to the social-emotional goals of Project Developmental Continuity. The system focuses on aspects of an individual child's behavior, verbal or nonverbal, that reflect the child's attitude toward himself, and on the child's social competence as demonstrated in his interaction with peers and adults.

Using a time sampling method, trained observers observe each child for five minutes at four different times during the day and code their behavior into four general categories: "noninvolved," "involved," "interacts with peer," and "interacts with adult." A fifth category, 'activity level," is included to provide information concerning the context in which these behaviors were observed. Each of these categories includes subcategories that are designed to identify the frequency and nature of specific behaviors within the general category.

Preschool Interpersonal Problem-Solving Test (Shure and Spivack, 1974)[1]. The PIPS attempts to assess the child's ability to name alternative solutions to a life-related problem--that of obtaining a toy from another child. Paper cutouts of boys, girls and toys are used in presenting the problem. Among inner city four-year-olds attending the Philadelphia Get Set day care program, those judged as better-adjusted by their teachers were able to conceptualize a greater number and a wider range of alternative solutions to real-life problems than were their more poorly adjusted classmates.

Pupil Observation Checklist (High/Scope Foundation, unpublished). This is a rating scale consisting of twelve 7-point bipolar adjectives derived from a similar scale used in the Home Start evaluation[2]. The tester rates each child using this instrument after he or she had administered all the other measures in the battery to the child. See Appendix H for details on the factor structure of this instrument.

PDC Faces Interview (High/Scope Foundation, unpublished). This test is designed to assess the child's attitude toward school and his teacher. The child is asked to point to one of 5 faces (which range from happy to sad) as he is asked questions about school and his teacher.

---

[1] Shure, M. B. & Spivack, G. The PIPS Test Manual. Philadelphia: Hahneman Medical College, 1974.

[2] Love, J. et al. National Home Start Evaluation Interim Report VII. Ypsilanti, MI: High/Scope Foundation, March 1976.

110

McCarthy Scales of Children's Abilities (McCarthy, 1972)[1]. These subtests consist of a series of tasks tapping problem-solving, psychomotor, and conceptual abilities, and are similar to the Wechsler scales, but with emphasis on age-related maturational indicators.

- Conceptual Grouping. The child's task is to manipulate and categorize objects along the dimensions of color, size, and shape according to verbally presented instructions.
- Verbal Memory. The child is asked to repeat sequences of words (Verbal Memory-1) and to repeat or retell as much as possible of a one paragraph story (Verbal Memory-3).
- Verbal Fluency. The child is asked to name as many members of specific categories (e.g., animals) as he/she can.
- Leg Coordination. Child performs motor tasks which involve the lower extremities, such as walking backwards or standing on one foot.
- Arm Coordination. Child bounces a rubber ball, catches a beanbag, and throws a beanbag through a hole in a target.
- Draw-A-Child. Child draws a picture of a child of the same sex.
- Opposite Analogies. The child is asked to supply the missing word in an analogy. For example, the sun is hot and ice is_____.

Wechsler Preschool and Primary Scale of Intelligence, Block Design subtest (Wechsler, 1967)[2]. The task requires reproducing (constructing) designs with flat colored blocks, either from the examiner's model or from a picture on a card. The measure taps problem-solving abilities, flexibility of response style, visual-motor organization, and execution.

CIRCUS subtest: Do You Know..? (Educational Testing Service, 1974)[3]. This is a general information test. The child chooses the picture which appropriately answers the examiner's question. This task taps the child's experience in a variety of areas (health, safety, social standards, consumer concepts).

---

[1] McCarthy, D. McCarthy Scales of Children's Abilities: Manual. New York Psychological Corporation, 1972.

[2] Wechsler, D. Wechsler Preschool and Primary Scale of Intelligence: Manual. New York: Psychological Corporation, 1967.

[3] Educational Testing Service. Circus Manual and Technical Report. Princeton: Educational Testing Service, 1974.

CIRCUS subtest: Say and Tell (Educational Testing Service, 1974)[1]. This test consists of two parts and taps children's language abilities. In the first part the child is given a pencil and asked attribute questions, e.g., "What color is it?"; in the second part the child is given two pennies and is asked to describe them. Scoring is based on categories of attributes which the child mentions.

Bilingual Syntax Measure (Burt, Dulay and Hernandez Ch., 1975)[2]. This test is designed to measure children's oral proficiency in English and/or Spanish grammatical structures. Simple questions are used with cartoon-type colored pictures to provide a conversational setting for eliciting natural speech. An analysis of the child's response yields a numerical indicator and a qualitative description of the child's structural language proficiency in standard English or standard Spanish. Responses are written down verbatim.

PDC Child Rating Scale (High/Scope Foundation, unpublished). This instrument is designed as a measure of social competence to be administered by the respective classroom teachers of the children rated. For each of the 39 items, specific behaviors such as "Uses words or wits to influence others" are rated on a 5-point scale according to frequency of occurrence ("Very frequently" to "Rarely"). See Appendix H for details on the factor structure of this instrument.

Adult Language Check. This measure is used in the bilingual/bicultural demonstration sites to obtain an indication of the languages the adults in the classroom use during their interactions with children. The interviewer sits in the classroom for a two hour period and records the language used by the teachers and aides approximately every five minutes.

Attrition, Handicap, and Attendance Information Sheet. To the extent that it was available, this information was collected for each child in the sample. In most cases it was obtained from the classroom teacher.

As Tables A-1 to A-3 show, not all the measures listed were administered at all the sites. The Spanish version of the Bilingual Syntax Measure was administered only to Spanish-speaking children, and the Adult Language Check was used only in classrooms where languages other than English were spoken by teachers, aides, or other adults. Due to the shortage of time available for training new testers, the Bilingual Syntax Measure was not administered in Georgia and the Classroom Observation System was not used in California or Michigan.

---

[1] Op. Cit.

[2] Burt, M., Dulay, H. & Hernandez, E. Bilingual Syntax Measure. New York: Harcourt, Brace, Jovanovich, 1975.

The PPLAT and the Child Rating Scale were administered in only a few sites since both instruments are still under development. Do You Know..?, Opposite Analogies, and WPPSI Block Design were given in a total of five sites, at the sites' options. (During the planning year each site had been allowed to select from a list of tests offered optionally for local inclusion.)

In Georgia, where a cross-sectional sample of elementary school children comprises the comparison group, only age-appropriate instruments (appropriateness indicated by test norms) were administered in the higher grades. And in Maryland, where a sample of third-graders was tested in order to examine the suitability of the instruments for children in the higher ranges of performance, certain tests were not included because it was already known that these were or were not suitable for children at this level.

Controlling for order effect in administering the BSM to bilingual children. Children who show facility in both Spanish and English receive both versions of the BSM. The order in which the two versions are administered is controlled so that during any single testing period half the children receive the Spanish version first and half receive the English version first. Further, the order is reversed with each successive testing so that, for example, a child who receives the Spanish version first in the fall would receive the English version first in the spring. The number of children who have actually received both versions is too small (18 in spring, 1976) to permit a statistical test of order effect, but the procedure described here should neutralize the effect if it indeed exists.

Table A-1

The Basic Battery of Measures

CHILD MEASURES	CALIFORNIA	COLORADO	CONNECTICUT	GEORGIA	FLORIDA	IOWA	MARYLAND	MICHIGAN	NEW JERSEY	TEXAS	UTAH	WASHINGTON	WEST VIRGINIA
Arm Coordination (McCarthy Scales)	X	X	X	X	X	X	X	X	X	X	X	X	X
Bilingual Syntax Measure (English Version)	X	X	X		X	X	X	X	X	X	X	X	X
Bilingual Syntax Measure (Spanish Version)	X[1]									X[1]			
Conceptual Grouping (McCarthy Scales)	X	X	X	X	X	X	X	X	X	X	X	X	X
Draw-A-Child (McCarthy Scales)	X	X	X	X	X	X	X	X	X	X	X	X	X
Leg Coordination (McCarthy Scales)	X	X	X	X	X	X	X	X	X	X	X	X	X
PDC Faces Interview (High/Scope)	X	X	X	X	X	X	X	X	X	X	X	X	X
Preschool Interpersonal Problem-Solving Test (PIPS)	X	X	X	X	X	X	X	X	X	X	X	X	X
Say and Tell (CIRCUS)	X	X	X	X	X	X	X	X	X	X	X	X	X
Verbal Fluency (McCarthy Scales)	X	X	X	X	X	X	X	X	X	X	X	X	X
Verbal Memory (McCarthy Scales)	X	X	X	X	X	X	X	X	X	X	X	X	X

OTHER MEASURES

OTHER MEASURES	CALIFORNIA	COLORADO	CONNECTICUT	GEORGIA	FLORIDA	IOWA	MARYLAND	MICHIGAN	NEW JERSEY	TEXAS	UTAH	WASHINGTON	WEST VIRGINIA
Adult Language Check (High/Scope)	X	X								X			
Attrition/Handicap/Attendance Information Sheet	X	X	X	X	X	X	X	X	X	X	X	X	X
Demographic Information Sheet (where applicable)	X	X	X	X	X	X	X	X	X	X	X	X	X
PDC Classroom Observation System (High/Scope)		X	X	X	X	X	X		X	X	X	X	X
Pupil Observation Checklist (POCL) (High/Scope)	X	X	X	X	X	X	X	X	X	X	X	X	X

[1] Administered to Spanish-speaking children in these sites.

Table A-2

Site-Selected and Experimental Measures

Child Measures	CALIFORNIA	COLORADO	CONNECTICUT	GEORGIA	FLORIDA	IOWA	MARYLAND	MICHIGAN	NEW JERSEY	TEXAS	UTAH	WASHINGTON	WEST VIRGINIA
Do You Know..? (CIRCUS)					X								X
Opposite Analogies (McCarthy Scales)		X									X		
PDC Child Rating Scale (High/Scope)		X	X	X		X					X		X
Preschool Productive Language Assessment Task (PPLAT--High/Scope)			X			X							
Block Design (WPPSI)								X					

Table A-3

Measures Administered in Georgia and Maryland

	GEORGIA					MARYLAND
CHILD MEASURES	Head Start	Kinder-garten	Grade 1	Grade 2	Grade 3	Grade 3
Conceptual Grouping	X	X	X	X	X	X
Verbal Fluency	X	X	X	X	X	X
Arm Coordination	X	X	X	X	X	X
Draw-A-Child	X	X	X	X	X	X
PDC Faces Interview	X	X	X	X	X	X
BSM-English						X
Say and Tell	X					
PIPS	X	X	X	X	X	X
Leg Coordination	X	X				
Verbal Memory	X	X	X	X	X	X
OTHER MEASURES						
PDC Classroom Observation System	X			X		
Pupil Observation Checklist (POCL)	X	X	X	X	X	X
Attrition/Handicap/Attendance Information Sheet	X	X	X	X	X	

APPENDIX B

Forms for Weekly Tester Monitoring

The forms reproduced here were used weekly by testers for mutual monitoring. The completed forms were returned regularly to High/Scope for continuing analysis. In this appendix, the categories beside which an X appears are those in which testers, as a group, made more errors than expected or than was judged tolerable.

Table B-1

ARM COORDINATION

Monitoring Form

Interviewer ______________________ Date ______________

Child's Name ______________________

INSTRUCTIONS: This form will provide High/Scope Foundation with information on how similar the interview administrations are within each site and across sites. The interviews must be administered in a standard or uniform way to insure comparability of the data. When you monitor another interviewer you should be recording the child's responses in your interview booklet and be watching for and noting whether any of the following errors occur during each of the interviews. You will fill out one of these monitoring forms for each interview you monitor.

Interview Administration Errors	Check if Occurred
1. Fails to have CORRECT INTERVIEWING MATERIALS; e.g., didn't have ball, beanbag, tape, etc.	
2. INCORRECT PLACEMENT of interview materials; e.g., didn't have target 6' from child, didn't kneel or bend when throwing beanbag to child, etc.	
3. INCORRECT WORDING of interview questions; e.g., doesn't follow the words in the interview booklet.	X
4. SKIPPED AN ITEM.	
5. SKIPPED A SECOND TRIAL, or gave a second trial when it should not have been given.	
6. STOPPED INTERVIEW INCORRECTLY; e.g., didn't give entire interview.	
7. REPEATS; repeated the interview question more than one time.	
8. ENCOURAGEMENTS; gave more than one encouragement per initial question and repeat; didn't give an encouragement when needed.	
9. SCORING; scored child's response incorrectly.	X
10. OTHER: (specify) ______________________ ______________________ ______________________	

Rapport with child (circle one): Poor Adequate Good

Name of Monitor ______________________

Table B-2

BILINGUAL SYNTAX MEASURE

Monitoring Form

Interviewer ______________________ Date ______________

Child's Name ______________________

INSTRUCTIONS: This form will provide High/Scope Foundation with information on how similar the interview administrations are within each site and across sites. The interviews must be administered in a standard or uniform way to insure comparability of the data. When you monitor another interviewer you should be recording the child's responses in your interview booklet and be watching for and noting whether any of the following errors occur during each of the interviews. You will fill out one of these monitoring forms for each interview you monitor.

Interview Administration Errors	Check if Occurred
1. Fails to have CORRECT INTERVIEWING MATERIALS; e.g., is missing the warm-up picture.	______
2. INCORRECT PLACEMENT of interview materials; e.g., doesn't place warm-up picture directly in front of child, doesn't place picture booklet directly in front of child; didn't put warm-up picture out of child's sight when using booklet, etc.	______
3. INCORRECT WORDING of interview questions; e.g., doesn't follow the words in the interview booklet, adds too many additional comments or questions.	______
4. SKIPPED AN ITEM.	______
5. STOPPED INTERVIEW INCORRECTLY; e.g., didn't stop after item 5 when child responded to only two of the first five items; stopped after three DK-R-NR instead of after four.	______
6. INCORRECT TIMING; e.g., didn't mark time started and time stopped on cover of interview booklet.	______
7. REPEATS; repeated the interview question more than one time; repeated the child's response verbally.	______
8. ENCOURAGEMENTS; gave more than one encouragement after the initial question; gave more than one encouragement after the repeat or didn't give an encouragement when it should have been given.	______
9. SCORING; not writing child's response exactly as said; not writing legibly; not indicating child's physical gestures, e.g., pointing, etc.	X
10. DEFINES WORDS; defining words for child during the non-preliminary questions.	______
11. OTHER: (specify) ______________________ ______________________	

Rapport with child (circle one): Poor Adequate Good

Table B-3

CONCEPTUAL GROUPING

Monitoring Form

Interviewer ______________________________ Date ________________

Child's Name ______________________________

INSTRUCTIONS: This form will provide High/Scope Foundation with information on how similar the interview administrations are within each site and across sites. The interviews must be administered in a standard or uniform way to insure comparability of the data. When you monitor another interviewer you should be recording the child's responses in your interview booklet and be watching for and noting whether any of the following errors occur during each of the interviews. You will fill out one of these monitoring forms for each interview you monitor.

Interview Administration Errors	Check if Occurred
1. Fails to have CORRECT INTERVIEWING MATERIALS, e.g., didn't have all of squares and circles.	____
2. INCORRECT PLACEMENT of interview materials, e.g., puts wrong materials on cardboard.	____
3. INCORRECT WORDING of interview questions, e.g., doesn't follow the words in the interview booklet.	____
4. SKIPPED AN ITEM.	____
5. STOPPED INTERVIEW INCORRECTLY, e.g., stopped after 3 failures or DK-R-NR instead of after 4.	____
6. INCORRECT TIMING, e.g., didn't record time started.	____
7. REPEATS; repeated the interview question more than one time.	____
8. SCORING; scored child's response incorrectly.	____
9. ENCOURAGEMENTS; gave more than one encouragement each time question was asked or didn't give encouragement when needed.	____
10. Other: (specify) ______________________________ ______________________________	

Rapport with child (circle one): Poor Adequate Good

Name of Monitor ____________________

Table B-4
DO YOU KNOW...?

Monitoring Form

Interviewer______________________________ Date____________________

Child's Name______________________________

INSTRUCTIONS: This form will provide High/Scope Foundation with information on how similar the interview administrations are within each site and across sites. The interviews must be administered in a standard or uniform way to insure comparability of the data. When you monitor another interviewer you should be recording the child's responses in your interview booklet and be watching for and noting whether any of the following errors occur during each of the interviews. You will fill out one of these monitoring forms for each interview you monitor.

Interview Administration Errors	Check if Occurred
1. Fails to have CORRECT INTERVIEWING MATERIALS, e.g., did not have manual and/or child's copy.	________
2. INCORRECT PLACEMENT of interview materials, e.g., did not place child's test booklet properly.	________
3. INCORRECT WORDING of interview questions, e.g., doesn't follow the words in the interview booklet.	________
4. SKIPPED AN ITEM.	________
5. STOPPED INTERVIEW INCORRECTLY, e.g., didn't test after 4 consecutive DK-R-NR's.	________
6. INCORRECT TIMING, e.g., didn't record time started and time stopped.	________
7. REPEATS; repeated the interview question more than one time.	________
8. SCORING; scored child's response incorrectly.	________
9. ENCOURAGEMENTS; gave more than one encouragement per initial question and per repeat or didn't give encouragement when needed.	________
10. Probes; did not ask, "Which one is your answer?", when child indicated two responses.	________
11. OTHER: (specify)________________________ ________________________	

Rapport with child (circle one): Poor Adequate Good

Name of Monitor______________________________

Table B-5
DRAW-A-CHILD
Monitoring Form

Interviewer ______________________________ Date ________________

Child's Name ______________________________

INSTRUCTIONS: This form will provide High/Scope Foundation with information on how similar the interview administrations are within each site and across sites. The interviews must be administered in a standard or uniform way to insure comparability of the data. When you monitor another interviewer you should be recording the child's responses in your interview booklet and be watching for and noting whether any of the following errors occur during each of the interviews. You will fill out one of these monitoring forms for each interview you monitor.

Interview Administration Errors	Check if Occurred
1. INCORRECT PLACEMENT of interview materials; e.g., didn't place blank page width-wise in front of child.	________
2. INCORRECT WORDING of interview questions; e.g., doesn't follow the words in the interview booklet.	________
3. INCORRECT TIMING; e.g., didn't mark time stopped.	________
4. REPEATS; repeated the interview question more than one time.	________
5. ENCOURAGEMENTS; failed to give one encouragement specified in interview booklet.	________
6. OTHER: (specify) ____________________ ______________________________ ______________________________	

Rapport with child (circle one): Poor Adequate Good

Name of Monitor ________________________

Table B-6

LEG COORDINATION

Monitoring Form

Interviewer ______________________________ Date ________________

Child's Name ______________________________

INSTRUCTIONS: This form will provide High/Scope Foundation with information on how similar the interview administrations are within each site and across sites. The interviews must be administered in a standard or uniform way to insure comparability of the data. When you monitor another interviewer you should be recording the child's responses in your interview booklet and be watching for and noting whether any of the following errors occur during each of the interviews. You will fill out one of these monitoring forms for each interview you monitor.

Interview Administration Errors	Check if Occurred
1. Fails to have CORRECT INTERVIEWING MATERIALS; e.g., didn't have 9-foot tape.	________
2. INCORRECT WORDING of interview questions; e.g., doesn't follow the words in the interview booklet.	________
3. SKIPPED AN ITEM.	________
4. SKIPPED A SECOND TRIAL or gave a second trial when it should not have been given.	________
5. STOPPED INTERVIEW INCORRECTLY; e.g., didn't stop after four DK-R-NR.	________
6. INCORRECT TIMING; e.g., didn't time child's standing on one foot correctly (interviewer and monitor should be within three seconds of each other), didn't mark time stopped.	________
7. REPEATS; gave a second demonstration when it should not have been given.	________
8. SCORING; scored child's response incorrectly.	________
9. ENCOURAGEMENTS; gave more than one encouragement on the initial question and more than one encouragement on the second trial or gave no encouragements when one should have been given.	________
10. OTHER: (specify) ______________________ ______________________________ ______________________________	

Rapport with child (circle one): Poor Adequate Good

Name of Monitor ______________________________

Table B-7

OPPOSITE ANALOGIES

Monitoring Form

Interviewer________________________________ Date____________________

Child's Name______________________________

INSTRUCTIONS: This form will provide High/Scope Foundation with information on how similar the interview administrations are within each site and across sites. The interviews must be administered in a standard or uniform way to insure comparability of the data. When you monitor another interviewer you should be recording the child's responses in your interview booklet and be watching for and noting whether any of the following errors occur during each of the interviews. You will fill out one of these monitoring forms for each interview you monitor.

Interview Administration Errors	Check When Occurs
1. INCORRECT WORDING of interview questions; e.g., doesn't follow the words in the interview booklet.	________
2. SKIPPED AN ITEM.	________
3. STOPPED INTERVIEW INCORRECTLY, e.g., doesn't stop interview after child failed items 1 and 2, or didn't stop after 3 DK-R-NR's.	________
4. INCORRECT TIMING; e.g., didn't record time started and time stopped.	________
5. REPEATS; repeated the interview question once. NO REPEATS ALLOWED.	________
6. SCORING; did not record the child's response correctly.	________
7. ENCOURAGEMENTS; gave more than one encouragement or didn't give encouragement when needed.	________
8. OTHER: (specify)____________________ ________________________________ ________________________________	

Rapport with child (circle one): Poor Adequate Good

Name of Monitor____________________________

Table B-8

PDC FACES INTERVIEW

Monitoring Form

Interviewer ______________________ Date ______________

Child's Name ______________________

INSTRUCTIONS: This form will provide High/Scope Foundation with information on how similar the interview administrations are within each site and across sites. The interviews must be administered in a standard or uniform way to insure comparability of the data. When you monitor another interviewer you should be recording the child's responses in your interview booklet and be watching for and noting whether any of the following errors occur during each of the interviews. You will fill out one of these monitoring forms for each interview you monitor.

Interview Administration Errors	Check Each Time Occurred
1. INCORRECT PLACEMENT of interview materials; e.g., didn't place faces page toward child.	______
2. INCORRECT WORDING of test questions; e.g., didn't follow the words in the interview booklet.	______
3. SKIPPED AN ITEM.	______
4. STOPPED INTERVIEW INCORRECTLY; e.g., didn't stop test after 4 consecutive DK-R-NR's.	______
5. INCORRECT TIMING; e.g., didn't mark time started and time stopped.	______
6. REPEATS; repeated the test question more than one time.	______
7. SCORING; didn't record the child's response correctly.	______
8. ENCOURAGEMENTS; gave more than one encouragement after the initial question or after the repeated question, or failed to give an encouragement when should have.	______
9. Proceeded with the interview when it was apparent that the child does not understand the faces and the concept behind them (happy-sad).	______
10. OTHER: (specify) ______________ ______________ ______________	

Rapport with child (circle one): Poor Adequate Good

Name of Monitor ______________________

127

Table B-9

PIPS

Monitoring Form

Interviewer ______________________________ **Date** ________________

Child's Name ______________________________

INSTRUCTIONS: This form will provide High/Scope Foundation with information on how similar the interview administrations are within each site and across sites. The interviews must be administered in a standard or uniform way to insure comparability of the data. When you monitor another interviewer you should be recording the child's responses in your interview booklet and be watching for and noting whether any of the following errors occur during each of the interviews. You will fill out one of these monitoring forms for each interview you monitor.

Interview Administration Errors	Check if Occurred
1. Fails to have CORRECT INTERVIEWING MATERIALS; e.g., missing one of the PIPS cutouts, etc.	________
2. INCORRECT PLACEMENT of interview materials; e.g., putting toy on wrong cut-out, placing cut-outs on table rather than on some kind of stand.	________
3. INCORRECT WORDING of interview questions; e.g., doesn't follow the words in the interview booklet.	X
4. SKIPPED AN ITEM.	________
5. STOPPED INTERVIEW INCORRECTLY; e.g., didn't stop interview after two consecutive stories in which child gave repetition of answers or DK-R-NR.	________
6. INCORRECT TIMING; didn't mark time started and time stopped.	________
7. PROBING, too many or too few; e.g., didn't probe when response required it or probed when child's answer was acceptable.	________
8. SCORING; recorded child's response incorrectly or failed to put child's response in correct response box.	________
9. OTHER: (specify) ______________________ ______________________ ______________________	

Rapport with child (circle one): Poor Adequate Good

Name of Monitor ______________________________

Table B-10

SAY AND TELL

Monitoring Form

Interviewer________________________________ Date__________________

Child's Name________________________________

INSTRUCTIONS: This form will provide High/Scope Foundation with information on how similar the interview administrations are within each site and across sites. The interviews must be administered in a standard or uniform way to insure comparability of the data. When you monitor another interviewer you should be recording the child's responses in your interview booklet and be watching for and noting whether any of the following errors occur during each of the interviews. You will fill out one of these monitoring forms for each interview you monitor.

Interview Administration Errors	Check if Occurred
1. Fails to have CORRECT INTERVIEW MATERIALS, e.g., used 2 nickels instead of pennies, used a red pencil, etc.	____
2. INCORRECT PLACEMENT of interview materials, e.g., didn't place pennies in child's hand, put pencil out of child's reach.	____
3. INCORRECT WORDING of interview questions, e.g., doesn't follow the words in the interview booklet.	____
4. SKIPPED AN ITEM.	____
5. STOPPED INTERVIEW INCORRECTLY, e.g., didn't give entire interview.	____
6. INCORRECT TIMING, e.g., didn't record time started and time stopped.	____
7. REPEATS; repeated the interview question more than one time.	____
8. ENCOURAGEMENTS; gave more than one encouragement or failed to give an encouragement.	____
9. SCORING; did not write child's response exactly as said, did not write legibly.	____
10. Other: (specify) ________________________________ ________________________________	

Rapport with child (circle one): Poor Adequate Good

Name of Monitor__________________________

Table B-11

VERBAL FLUENCY

Monitoring Form

Interviewer ______________________ Date ______________

Child's Name ______________________

INSTRUCTIONS: This form will provide High/Scope Foundation with information on how similar the interview administrations are within each site and across sites. The interviews must be administered in a standard or uniform way to insure comparability of the data. When you monitor another interviewer you should be recording the child's responses in your interview booklet and be watching for and noting whether any of the following errors occur during each of the interviews. You will fill out one of these monitoring forms for each interview you monitor.

Interview Administration Errors	Check if Occurred
1. INCORRECT WORDING of interview questions; e.g., doesn't follow the words in the interview booklet.	______
2. SKIPPED AN ITEM.	______
3. STOPPED INTERVIEW INCORRECTLY; e.g., didn't give entire interview.	______
4. INCORRECT TIMING; e.g., allowed the child 26 seconds to name all the toys he could think of instead of 20 seconds.	______
5. REPEATS; repeated the interview question.	______
6. SCORING; didn't record child's response exactly as said, didn't write legibly.	______
7. ENCOURAGEMENTS; failed to say appropriate encouragement after 5 seconds, or encouraged too many times.	X
8. OTHER: (specify) ______________ ______________________ ______________________	

Rapport with child (circle one): Poor Adequate Go

Name of Monitor ______________________

Table B-12
VERBAL MEMORY
Monitoring Form

Interviewer ______________________________ Date ________________

Child's Name ______________________________

INSTRUCTIONS: This form will provide High/Scope Foundation with information on how similar the interview administrations are within each site and across sites. The interviews must be administered in a standard or uniform way to insure comparability of the data. When you monitor another interviewer you should be recording the child's responses in your interview booklet and be watching for and noting whether any of the following errors occur during each of the interviews. You will fill out one of these monitoring forms for each interview you monitor.

Interview Administration Errors	Check if Occurred
1. INCORRECT WORDING of interview questions; e.g., doesn't follow the words in the interview booklet.	________
2. SKIPPED AN ITEM.	________
3. STOPPED INTERVIEW INCORRECTLY; e.g., failed to give entire interview.	________
4. REPEATS; repeated the interview question.	________
5. SPEED; read the words too quickly for the child or allowed too much time between the words.	________
6. ENCOURAGEMENTS; encouraged the child more than once or didn't encourage the child at all when he didn't respond.	________
7. SCORING; failed to record child's response correctly or wrote child's response illegibly.	________
8. OTHER: (specify) ______________________________ ______________________________	

Rapport with child (circle one): Poor Adequate Good

Name of Monitor ______________________________

Table B-13

WPPSI BLOCK DESIGN
Monitoring Form

Interviewer ______________________________ Date ________________

Child's Name ______________________________

INSTRUCTIONS: This form will provide High/Scope Foundation with information on how similar the interview administrations are within each site and across sites. The interviews must be administered in a standard or uniform way to insure comparability of the data. When you monitor another interviewer you should be recording the child's responses in your interview booklet and be watching for and noting whether any of the following errors occur during each of the interviews. You will fill out one of these monitoring forms for each interview you monitor.

Interview Administration Errors	Check if Occurred
1. Fails to have CORRECT INTERVIEWING MATERIALS; e.g., doesn't have all 14 blocks, doesn't have picture booklet.	________
2. INCORRECT PLACEMENT of interview materials; e.g., makes incorrect WPPSI design, uses wrong blocks in making design, giving wrong blocks to child.	________
3. INCORRECT WORDING of interview questions; e.g., doesn't follow the words in the interview booklet.	________
4. SKIPPED AN ITEM.	________
5. SKIPPED A SECOND TRIAL, or gave a second trial when it should not have been given.	________
6. STOPPED INTERVIEW INCORRECTLY; e.g., didn't stop after two consecutive failures, or didn't give item 4 after child failed items 2 and 3.	________
7. INCORRECT TIMING; e.g., allowed the child more or less time to make the design than the instructions indicated (should be within 5 seconds of time limit), didn't mark time started and time stopped.	________
8. REPEATS; gave a demonstration when it should not have been given or failed to give a demonstration; repeated the interview question.	________
9. ENCOURAGEMENTS; gave more than one encouragement or none at all with the initial question and more than one encouragement or none at all on the second trial.	________
10. SCORING; scored child's response incorrectly.	________
11. ROTATIONS and GAPS; failed to correct child's rotations or ask child, "Is that right?" when he left more than 1/4 inch between his blocks.	________
12. OTHER: (specify) ______________________________ ______________________________	

Rapport with child (circle one): Poor Adequate Good

Name of Monitor: ______________________________

APPENDIX C

Flow Charts for the Analysis Procedure

Figure C-1
Flow Chart for Step 1: Does the Spring 1976 Internal Consistency Coefficient Indicate Adequate Reliability?

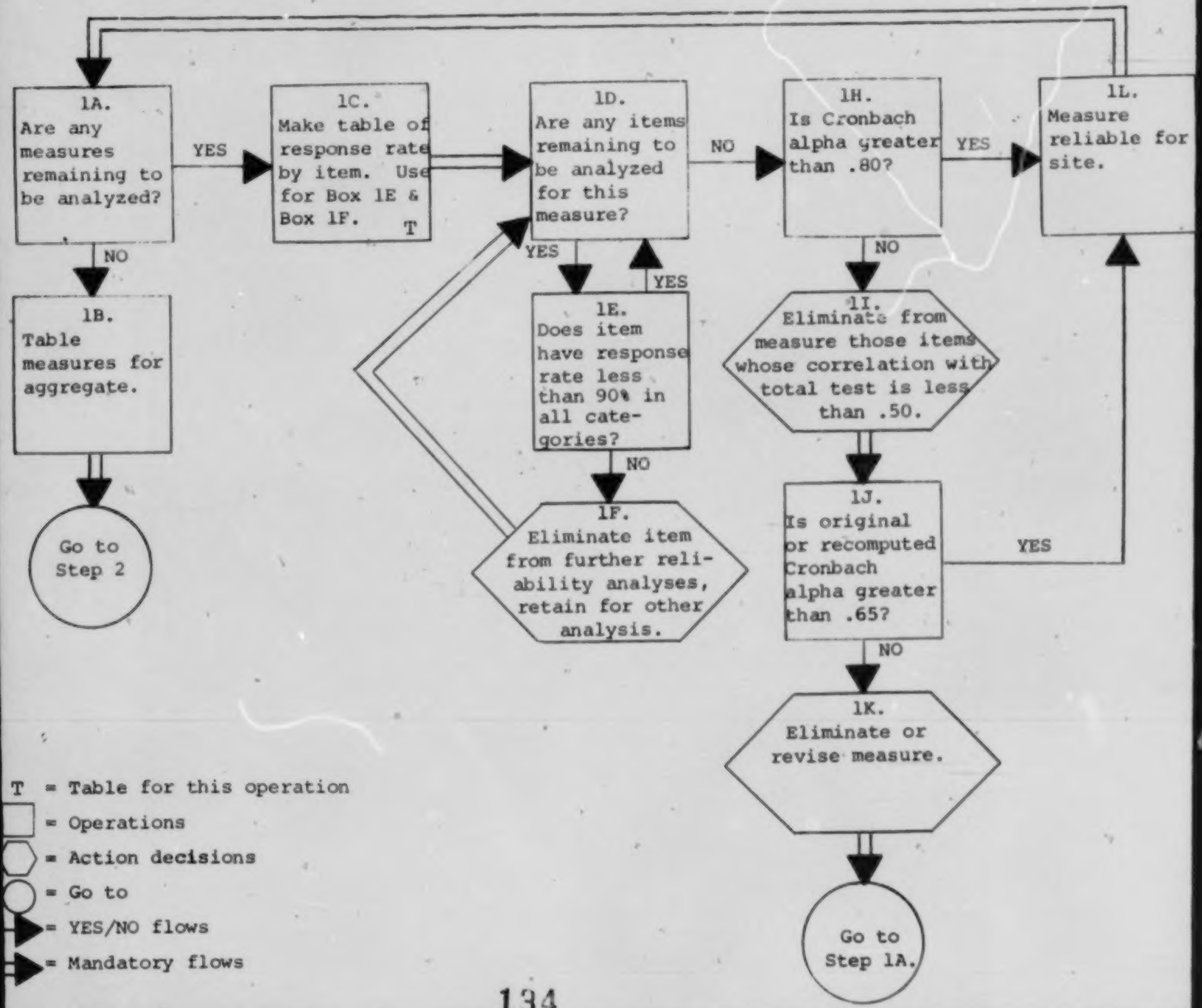
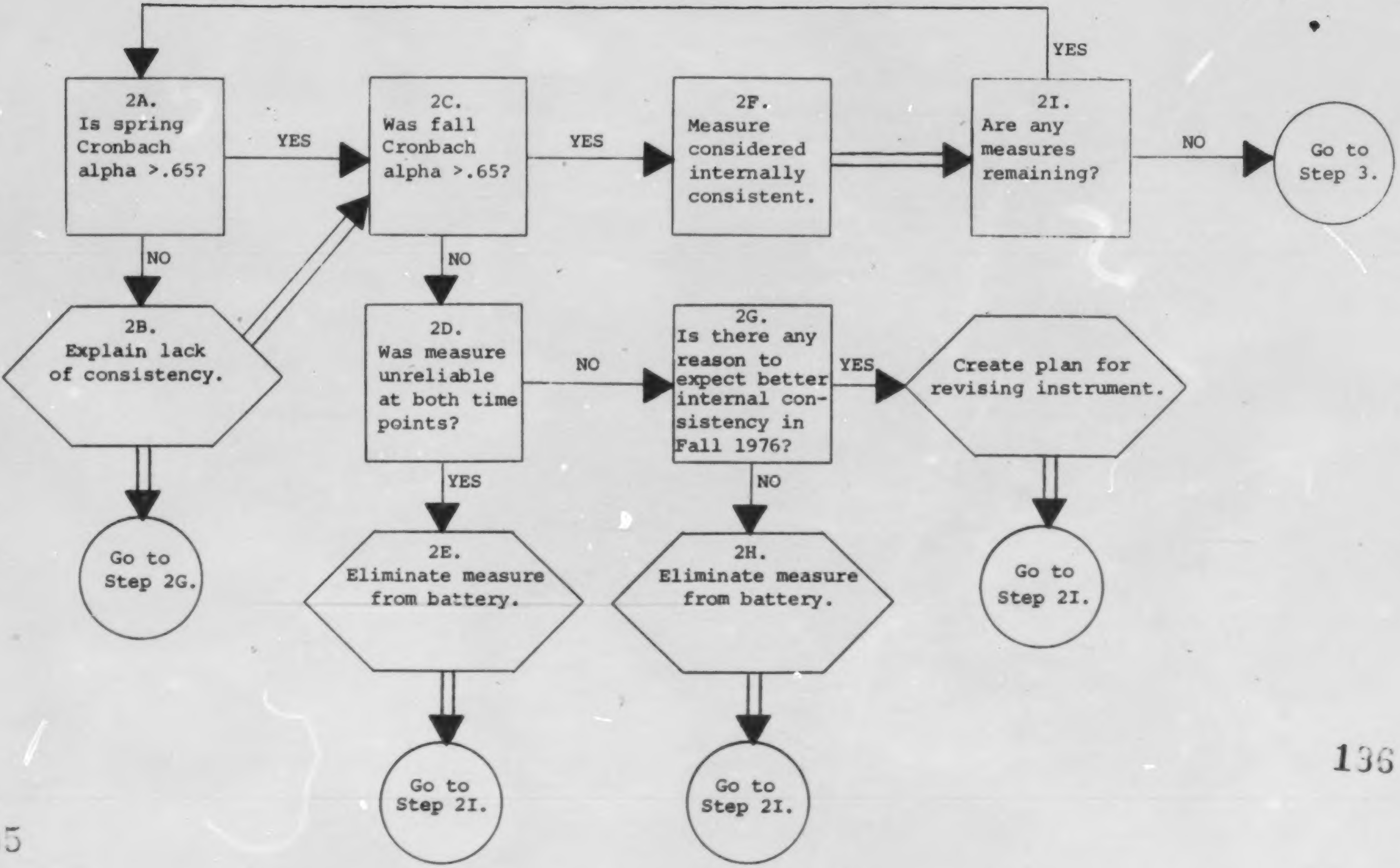

1A. Are any measures remaining to be analyzed?
YES
NO
1B. Table measures for aggregate.
Go to Step 2
1C. Make table of response rate by item. Use for Box 1E & Box 1F. T
1D. Are any items remaining to be analyzed for this measure?
YES
NO
1E. Does item have response rate less than 90% in all categories?
YES
NO
1F. Eliminate item from further reliability analyses, retain for other analysis.
1H. Is Cronbach alpha greater than .80?
YES
NO
1I. Eliminate from measure those items whose correlation with total test is less than .50.
1J. Is original or recomputed Cronbach alpha greater than .65?
YES
NO
1K. Eliminate or revise measure.
Go to Step 1A.
1L. Measure reliable for site.
T = Table for this operation
= Operations
= Action decisions
= Go to
= YES/NO flows
= Mandatory flows

Figure C-2

Flow Chart for Step 2: Is the Internal Consistency Coefficient Constant from Fall to Spring?



YES
2A.
Is spring Cronbach alpha >.65?
YES
2C.
Was fall Cronbach alpha >.65?
YES
2F.
Measure considered internally consistent.
2I.
Are any measures remaining?
NO
Go to Step 3.
NO
2B.
Explain lack of consistency.
NO
2D.
Was measure unreliable at both time points?
NO
2G.
Is there any reason to expect better internal consistency in Fall 1976?
YES
Create plan for revising instrument.
Go to Step 2G.
YES
2E.
Eliminate measure from battery.
NO
2H.
Eliminate measure from battery.
Go to Step 2I.
Go to Step 2I.
Go to Step 2I.

Figure C-3

Flow Chart for Step 3: Does the Test-Retest Coefficient Indicate Adequate Reliability?

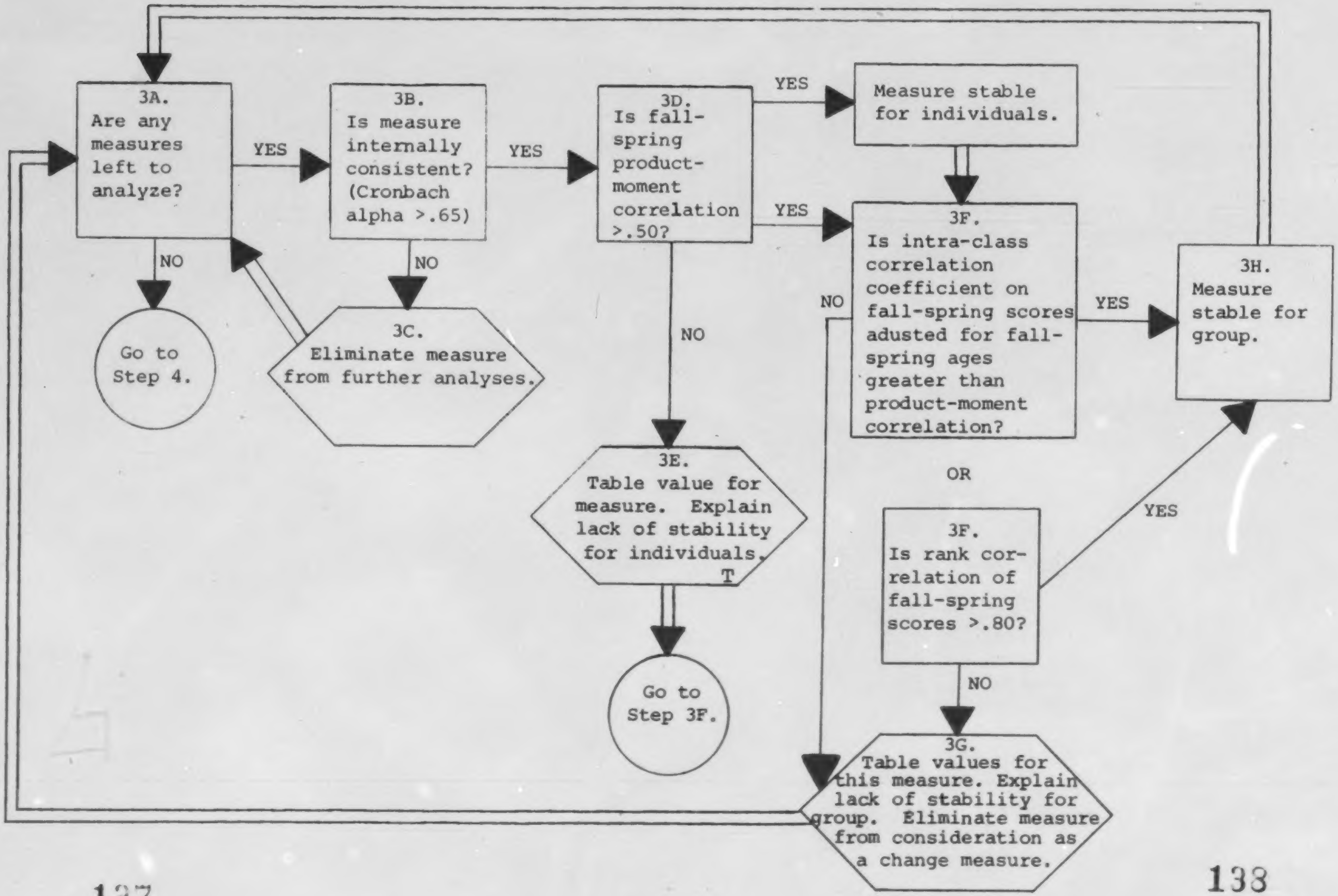

3A.
Are any measures left to analyze?
YES
NO
Go to Step 4.
3B.
Is measure internally consistent? (Cronbach alpha >.65)
YES
NO
3C.
Eliminate measure from further analyses.
3D.
Is fall-spring product-moment correlation >.50?
YES
Measure stable for individuals.
YES
NO
3E.
Table value for measure. Explain lack of stability for individuals.
Go to Step 3F.
3F.
Is intra-class correlation coefficient on fall-spring scores adusted for fall-spring ages greater than product-moment correlation?
NO
YES
3H.
Measure stable for group.
OR
3F.
Is rank correlation of fall-spring scores >.80?
YES
NO
3G.
Table values for this measure. Explain lack of stability for group. Eliminate measure from consideration as a change measure.

Figure C-4

Flow Chart for Step 4: Are the Measures Valid?

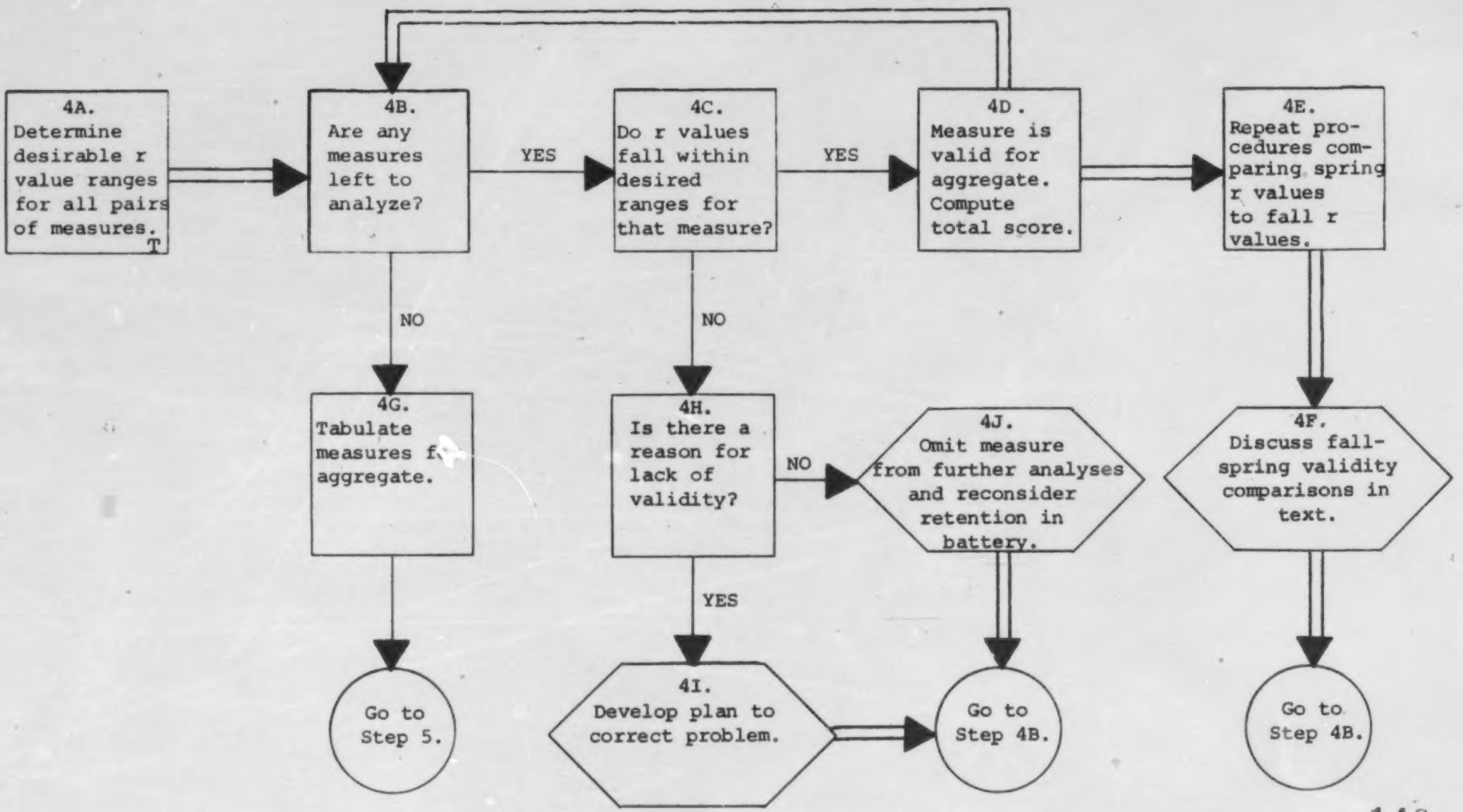

4A.
Determine desirable r value ranges for all pairs of measures.
4B.
Are any measures left to analyze?
YES
4C.
Do r values fall within desired ranges for that measure?
YES
4D.
Measure is valid for aggregate. Compute total score.
4E.
Repeat procedures comparing spring r values to fall r values.
NO
NO
4G.
Tabulate measures for aggregate.
4H.
Is there a reason for lack of validity?
NO
4J.
Omit measure from further analyses and reconsider retention in battery.
4F.
Discuss fall-spring validity comparisons in text.
YES
Go to Step 5.
4I.
Develop plan to correct problem.
Go to Step 4B.
Go to Step 4B.

Figure C-5

Flow Chart for Step 5: Are the Measures Sensitive to Fall-Spring Change?

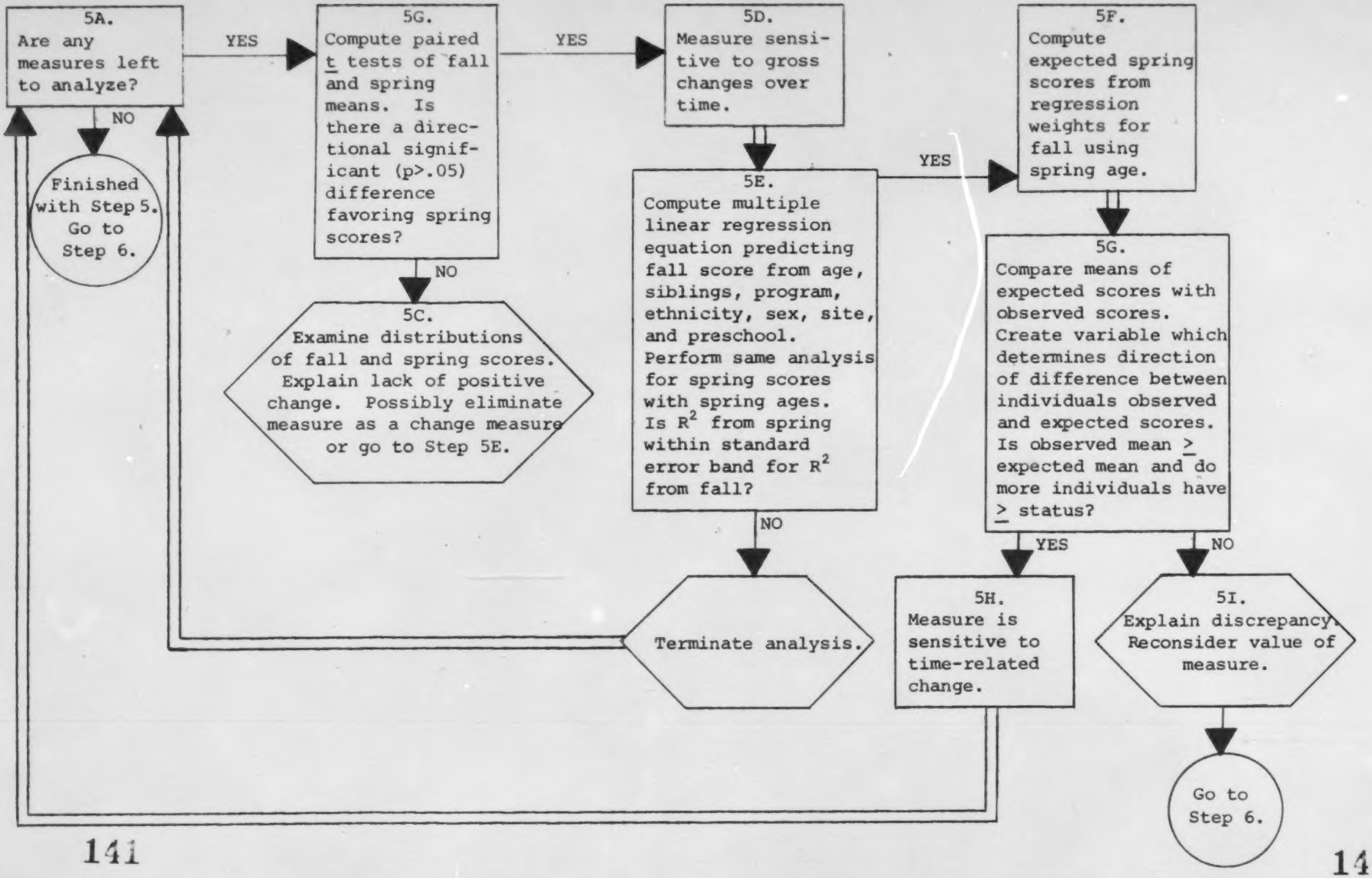
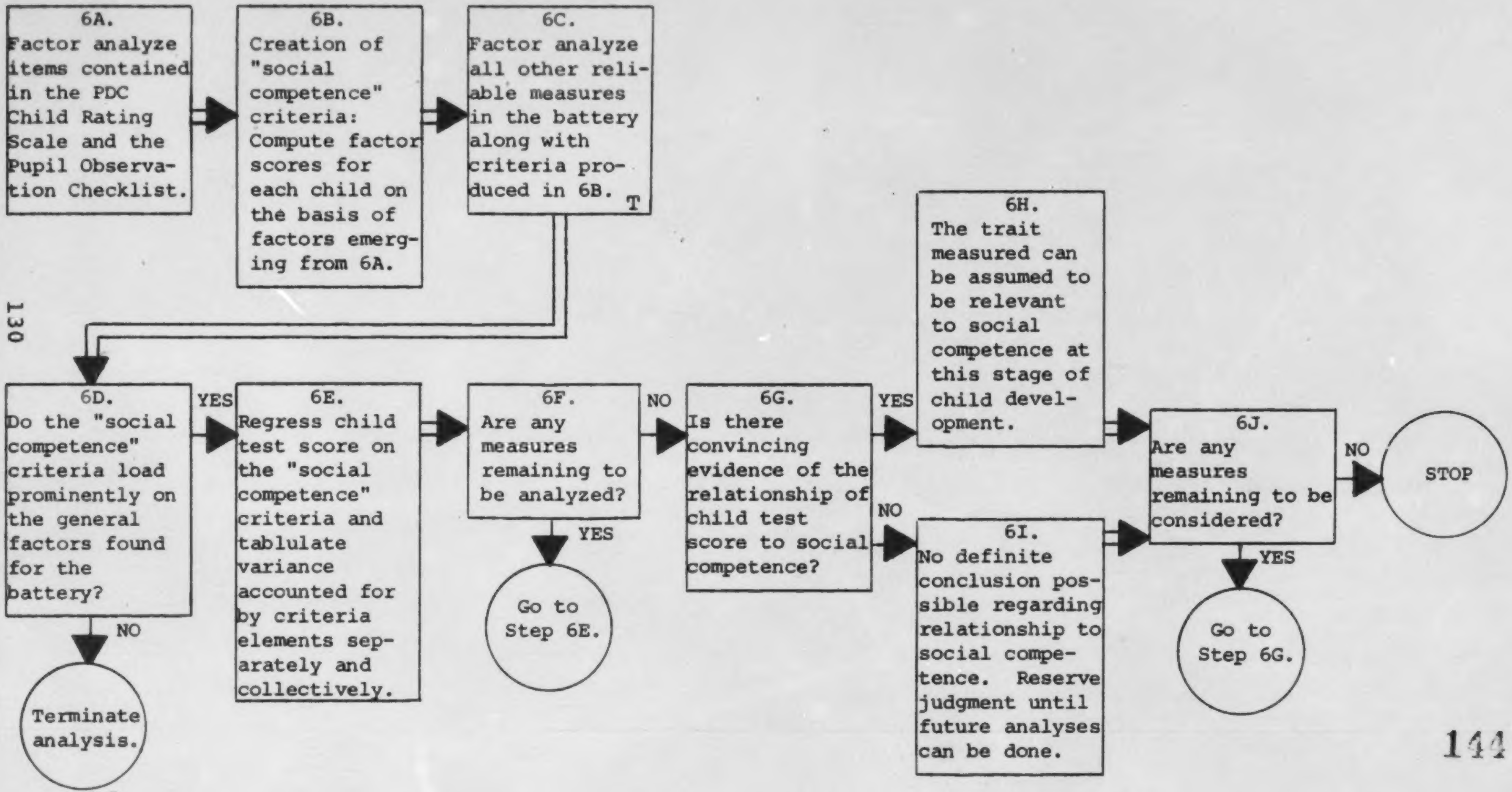

5A. Are any measures left to analyze?
YES
NO
Finished with Step 5. Go to Step 6.
5G. Compute paired t tests of fall and spring means. Is there a directional significant (p>.05) difference favoring spring scores?
YES
NO
5C. Examine distributions of fall and spring scores. Explain lack of positive change. Possibly eliminate measure as a change measure or go to Step 5E.
5D. Measure sensitive to gross changes over time.
5E. Compute multiple linear regression equation predicting fall score from age, siblings, program, ethnicity, sex, site, and preschool. Perform same analysis for spring scores with spring ages. Is R2 from spring within standard error band for R2 from fall?
YES
NO
Terminate analysis.
5F. Compute expected spring scores from regression weights for fall using spring age.
5G. Compare means of expected scores with observed scores. Create variable which determines direction of difference between individuals observed and expected scores. Is observed mean ≥ expected mean and do more individuals have ≥ status?
YES
NO
5H. Measure is sensitive to time-related change.
5I. Explain discrepancy. Reconsider value of measure.
Go to Step 6.

Figure C-6

Flow Chart for Step 6: Are the Measures Related to Social Competence?


6A.
Factor analyze items contained in the PDC Child Rating Scale and the Pupil Observation Checklist.
6B.
Creation of "social competence" criteria: Compute factor scores for each child on the basis of factors emerging from 6A.
6C.
Factor analyze all other reliable measures in the battery along with criteria produced in 6B.
T
6D.
Do the "social competence" criteria load prominently on the general factors found for the battery?
YES
NO
Terminate analysis.
6E.
Regress child test score on the "social competence" criteria and tablulate variance accounted for by criteria elements separately and collectively.
6F.
Are any measures remaining to be analyzed?
YES
Go to Step 6E.
NO
6G.
Is there convincing evidence of the relationship of child test score to social competence?
YES
6H.
The trait measured can be assumed to be relevant to social competence at this stage of child development.
NO
6I.
No definite conclusion possible regarding relationship to social competence. Reserve judgment until future analyses can be done.
6J.
Are any measures remaining to be considered?
NO
STOP
YES
Go to Step 6G.

APPENDIX D

Response Distributions for Child Measures
Spring 1976 Data

A. English-language measures, Head Start level (Tables D-1 through D-13)

B. Spanish-language measures, Head Start level (Tables D-14 through D-25)

C. English-language measures, kindergarten through Grade Three (Tables D-26 through D-70)

TABLE D-1

BILINGUAL SYNTAX MEASURE-ENGLISH (BSM-E)

RESPONSE DISTRIBUTION FOR: **ENGLISH-DOMINANT CHILDREN**

N= 430

ITEM NUMBER	PERCENT OF CORRECT RESPONSE	PERCENT NO RESPONSE
1 **(n=430)**	91.6	0.2
3 **(n=430)**	50.0	2.6
6 (n=421)	69.8	1.4
7 (n=421)	42.1	0.0
8 (n=421)	82.6	0.0
9 (n=421)	54.2	1.9
11 (n=421)	51.4	1.2
13 (n=421)	2.3	1.4
14 (n=421)	45.8	2.3
15 (n=421)	47.4	1.2
16 (n=421)	46.3	0.7
17 (n=421)	46.3	1.9
18 (n=420)	86.3	0.0
19 (n=420)	82.6	1.9
20 (n=419)	84.7	0.2
22 (n=419)	73.3	0.2
23 (n=419)	3.3	2.8
24 (n=419)	62.8	0.2

TABLE D-2

BILINGUAL SYNTAX MEASURE-ENGLISH (BSM-E)

RESPONSE DISTRIBUTION FOR: **SPANISH-DOMINANT CHILDREN**

N= 10

ITEM NUMBER	PERCENT OF CORRECT RESPONSE	PERCENT NO RESPONSE
1 **(n= 10)**	20.0	20.0
3 **(n= 10)**	10.0	50.0
6 (n= 3)	0.0	10.0
7 (n= 3)	0.0	0.0
8 (n= 3)	0.0	10.0
9 (n= 3)	10.0	0.0
11 (n= 2)	10.0	10.0
13 (n= 2)	0.0	0.0
14 (n= 2)	0.0	0.0
15 (n= 2)	10.0	0.0
16 (n= 2)	10.0	0.0
17 (n= 2)	0.0	0.0
18 (n= 2)	10.0	0.0
19 (n= 2)	10.0	0.0
20 (n= 2)	10.0	0.0
22 (n= 2)	0.0	0.0
23 (n= 2)	0.0	0.0
24 (n= 2)	0.0	0.0

TABLE D-3

CONCEPTUAL GROUPING (ENGLISH VERSION)

RESPONSE DISTRIBUTION FOR: **ENGLISH-DOMINANT CHILDREN**

N = 462

	PERCENT OF SUBJECTS SCORING:			
ITEM NUMBER	*TWO* POINTS	*ONE* POINT	*NO* POINTS	*NO* RESPONSE
1 **(n=462)**	0.0	98.5	0.6	0.0
2 **(n=462)**	0.0	85.7	13.4	0.0
3 **(n=462)**	0.0	93.5	5.6	0.0
4 **(n=462)**	71.4	5.8	21.9	0.0
5 (n=462)	33.3	10.6	55.2	0.0
6 (n=462)	0.0	36.6	62.6	0.0
7 (n=462)	0.0	39.8	59.3	0.0
8 (n=407)	0.0	37.9	48.5	0.9
9 (n=330)	4.8	37.7	27.1	1.1

TABLE D-4

SAY AND TELL (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: **ENGLISH-DOMINANT CHILDREN**

N= 447

	PERCENT GIVING CORRECT RESPONSE	PERCENT NO RESPONSE
I. PENCIL		
NAME	95.7	0.4
COLOR	83.9	0.7
SHAPE	38.9	12.3
MATERIAL	38.3	11.4
FUNCTION	94.4	0.7
ANOTHER OBJECT FOR FUNCTION	16.3	12.1
ADDITIONAL INFORMATION	49.4	17.9
II. PENNIES		
LABEL	85.9	0.2
CLASS	29.3	0.2
COLOR	2.2	0.2
SHAPE	11.9	0.2
MATERIAL	1.6	0.2
FUNCTION	69.4	0.2
NUMBER	13.9	0.2
OTHER PHYSICAL CHARACTARISTICS	21.5	0.2
COMPARATIVE CHARACTERISTICS	0.7	0.2

TABLE- D-5

VERBAL FLUENCY (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: **ENGLISH-DOMINANT CHILDREN**

N = 458

	NO. OF EXAMPLES		
	MEAN	STANDARD DEVIATION	PERCENT NO RESPONSE
I. ANIMALS	2.2	2.1	16.6
II. THINGS TO EAT	2.6	1.7	6.6
III. NAMES	2.3	1.8	11.4
IV. TOYS	2.0	1.5	8.1

TABLE-D-6

VERBAL MEMORY (MSCA) (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: **ENGLISH-DOMINANT CHILDREN**

N = 443

1. ITEM	NUMBER OF WORDS REPEATED 4	3	2	1	0	CORRECT SEQUENCE
1.TOY-CHAIR-LIGHT	0	80	11	2	1	89 %
2.DOLL-DARK-COAT	0	67	25	4	0	91 %
3.AFTER-COLOR-FUNNY -TODAY	62	22	5	3	1	82 %
4.AROUND-BECAUSE-UNDER -NEVER	29	36	20	6	1	77 %

	MC CARTHY SCORE 4	3	2	1	0	NO RESPONSE
1.TOY-CHAIR-LIGHT	0	77	12	3	1	6 %
2.DOLL-DARK-COAT	0	64	28	4	0	4 %
3.AFTER-COLOR-FUNNY -TODAY	55	28	6	4	1	7 %
4.AROUND-BECAUSE-UNDER -NEVER	25	33	23	9	1	8 %

3. STORY PART	CREDIT FOR PARTS OF THE STORY ONE POINT	TWO POINTS	NO RESPONSE
1.TERM USED FOR BOB	70.9%	7.7%	19.4%
2.TERM USED FOR WOMAN	58.2	20.3	19.4
3.TERM USED FOR LETTERS	67.3	11.3	19.4
4.BOB WALKING TO STORE	20.8	57.8	19.4
5.BOB SAW WOMAN	18.3	60.3	19.4
6.WIND BLEW LETTERS	38.6	40.0	19.4
7.BOB SHOUTED, "..."	14.0	64.6	19.4
8.BOB WAS CAREFUL	32.7	45.8	19.4
9.BOB PICKED UP LETTERS	49.7	28.9	19.4
10.WOMAN WAS HAPPY	21.2	57.3	19.4
11.WOMAN THANKED BOB	11.7	66.8	19.4

TABLE- D-7

ARM COORDINATION (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: **ENGLISH-DOMINANT CHILDREN**

N =457

	PERCENT RESPONSE	PERCENT NO RESPONSE
I. BALL BOUNCING		
SCORE		1.75
7	1.31	
6	0.44	
5	1.97	
4	5.03	
3	27.13	
2	32.17	
1	30.20	
0	0.0	
II. BEANBAG CATCH GAME		
A. TOTAL TIMES BAG CAUGHT		0.22
3	6.56	
2	14.66	
1	27.57	
0	50.98	
B. BAG CAUGHT PREFERRED HAND		0.22
3	1.09	
2	4.38	
1	11.82	
0(N= 456)	82.28	
C. BAG CAUGHT NON-PREFERRED HAND		0.22
3	1.31	
2	2.84	
1	5.47	
0(N= 454)	89.50	
III. BEANBAG GAME		
A. PREFERRED HAND		0.44
6	0.0	
5	0.88	
4	3.06	
3	8.32	
2	18.60	
1	28.67	
0	40.04	
B. NON-PREFERRED HAND		0.44
6	0.0	
5	0.22	
4	2.19	
3	3.94	
2	10.72	
1	17.94	
0	24.95	

TABLE D-8

DRAW-A-CHILD (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: **ENGLISH-DOMINANT CHILDREN**

N = 456

BODY PART	PERCENT OF CHILDREN HAVING BODY PART PRESENT IN DRAWING
HEAD	93.9
HAIR	47.4
EYES	91.0
NOSE	62.9
MOUTH	74.1
NECK	6.6
TRUNK	53.9
ARMS AND HANDS	52.2
ATTACHMENT OF ARMS	28.7
LEGS AND FEET	76.8
NO DRAWING	0.4

TABLE D-9

LEG COORDINATION(ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: ENGLISH-DOMINANT CHILDREN

N = 446

	PERCENT RESPONSE	PERCENT NO RESPONSE
1.WALKING BACKWARD		0.9
5 OR MORE STEPS(2 PTS)	94.8	
2-4 STEPS(1 PT)	0.0	
0 PTS	4.0	
2.WALKING ON TIPTOE		0.0
5 OR MORE STEPS(2 PTS)	1.3	
2-4 STEPS(1 PT)	0.9	
0 PTS	97.1	
3.WALKING STRAIGHT LINE		0.2
2PTS	92.4	
1 PT	6.3	
0 PTS	1.1	
4.STANDING ON ONE FOOT		1.8
10 SEC OR MORE(2 PTS)	43.7	
3-9 SEC(1 PT)	44.2	
0-2 SEC(0 PTS)	9.0	
5.STANDING ON OTHER FOOT		2.0
10 SEC OR MORE(2 PTS)	40.4	
3-9 SEC(1 PT)	45.3	
0-2 SEC(0 PTS)	10.8	
6.SKIPPING		5.4
3 PTS	68.4	
2 PTS	9.9	
1 PT	6.7	
0 PTS	9.4	

TABLE D-10

PUPIL OBSERVATION CHECKLIST (POCL, FORM DC)

RESPONSE DISTRIBUTION FOR: **ENGLISH-DOMINANT CHILDREN**

N = 462

	PERCENT OF CHILDREN IN RATINGS CATEGORIES							PERCENT NOT RATED
CHILD'S TESTING BEHAVIOR	7	6	5	4	3	2	1	
COOPERATIVE	6.5	11.5	19.5	35.3	12.3	4.1	2.4	8.4
SOCIABLE	5.8	10.4	18.2	32.5	17.3	5.2	2.2	8.4
INVOLVED	3.5	10.8	16.0	42.6	13.0	4.5	0.9	8.7
TALKATIVE	8.4	8.9	17.1	32.9	16.2	6.1	1.7	8.7
ATTENTIVE	2.2	9.5	15.4	37.9	13.9	7.4	4.5	9.3
ACTIVE	5.2	10.6	14.3	47.6	11.3	1.7	0.2	9.1
RELAXED	4.3	9.3	15.8	50.6	8.4	1.9	0.6	8.9
QUICK TO RESPOND	2.2	10.0	14.1	32.3	23.6	5.4	3.9	8.7
ATTEMPTS DIFFICULT TASKS	3.5	8.7	12.3	40.3	17.1	6.7	2.8	8.7
KEEPS TRYING	2.8	9.3	14.7	39.0	16.5	6.1	3.2	8.4
SELF-CONFIDENT	1.3	8.0	12.6	40.3	14.7	6.5	3.7	13.0
HANDICAP AFFECTED PERFORMANCE	0.6	0.2	0.6	0.4	0.6	0.4	0.9	96.1

TABLE- D-11

FDC CHILD RATING SCALE (ENGLISH ONLY RATED BY TEACHER)
RESPONSE DISTRIBUTION FOR: ENGLISH-DOMINANT CHILDREN
TOTAL NUMBER OF RATINGS: 270

LENGTH OF TIME TEACHER HAD KNOWN CHILD:	PERCENT RESPONDING
LESS THAN MONTH	0.0
ONE-SIX MONTHS	0.0
SIX-TWELVE MONTHS	27.8
MORE THAN 12 MONTHS	1.9

HOW WELL TEACHER KNOWS CHILD	
NOT WELL	0.4
AVERAGE	20.0
BETTER THAN AVERAGE	7.4

I. SOCIAL COMPETENCE

	FREQUENTLY	MORE THAN AVERAGE	AVERAGE	LESS THAN AVERAGE	RARELY	NO RESPONSE
1.	4.8	6.3	13.3	4.4	0.4	0.0
2.	1.5	8.5	11.1	5.2	3.3	0.0
3.	6.7	7.4	14.8	5.6	1.1	0.0
4.	0.4	5.6	18.1	3.7	1.9	0.0
5.	0.0	4.8	17.4	4.4	3.0	0.0
6.	2.2	8.1	15.2	3.3	0.7	0.0
7.	3.3	7.0	16.7	2.2	0.0	0.0
8.	4.8	5.2	14.1	4.1	1.1	0.0
9.	3.0	4.4	17.0	4.8	0.4	0.0
10.	3.0	4.8	16.3	4.4	1.1	0.0
11.	1.5	5.2	13.3	7.0	2.6	0.0
12.	1.9	6.3	13.7	4.1	3.3	0.0
13.	1.1	6.3	12.2	7.4	2.6	0.0
14.	4.1	7.4	15.6	2.6	0.0	0.0
15.	1.9	5.9	18.1	1.5	0.0	0.0
16.	6.3	8.9	13.7	0.7	0.0	0.0
17.	3.3	7.4	16.3	2.2	0.4	0.0
18.	2.2	7.0	17.8	2.2	0.4	0.0
19.	4.4	5.6	15.9	3.0	0.7	0.0
20.	3.3	5.2	19.6	1.1	0.4	0.0
21.	1.9	3.0	20.4	4.1	0.4	0.0
22.	2.2	4.4	20.4	2.2	0.4	0.0
23.	4.8	7.4	15.6	1.5	0.4	0.0
24.	4.4	8.1	14.4	1.9	0.7	0.0

II. LEARNING HOW TO LEARN

	FREQUENTLY	MORE THAN AVERAGE	AVERAGE	LESS THAN AVERAGE	RARELY	NO RESPONSE
25.	3.1	8.5	15.6	1.5	0.0	0.0
26.	2.2	9.6	15.2	2.6	0.0	0.0
27.	0.7	4.8	17.4	4.8	1.5	0.0
28.	0.7	3.0	23.0	1.9	1.1	0.0
29.	0.0	2.2	15.2	6.7	5.6	0.0
30.	0.4	5.2	16.7	5.9	1.5	0.0
31.	0.4	4.4	17.8	3.7	3.3	0.0
32.	2.2	4.8	17.8	4.4	0.4	0.0
33.	5.2	8.1	16.3	0.0	0.0	0.0
34.	5.6	10.7	12.2	1.1	0.0	0.0
35.	0.0	3.3	21.5	4.1	0.7	0.0
36.	0.0	1.5	21.5	4.8	1.9	0.0
37.	2.2	6.7	16.7	4.1	0.0	0.0
38.	3.0	7.0	16.7	3.0	0.0	0.0
39.	43.3	5.2	0.0	3.0	0.0	7.4

TABLE D-12

FACES

RESPONSE DISTRIBUTION FOR: ENGLISH-DOMINANT CHILDREN

Variables are defined as follows:

801. Feelings about eating candy
802. Feelings about falling down
8801. Feelings about school this year
8802. Feelings about school next year
8803. Feelings about teacher
8804. Teachers feelings about child

```
     VARIABLE          N  MINIMUM      MAXIMUM       MEAN      STD DEV

  801.CANDY          437   1.0000       5.0000      3.7941      1.2336

  802.FALL           434   1.0000       5.0000      2.3779      1.2139

 8801.FSCHNOW        155   1.0000       5.0000      3.4903     .1.1360

 8802.FSCHNXY        155   1.0000       5.0000      3.0774      1.1927

 8803.FCH.TO.T       155   1.0000       5.0000      3.5102      .97485

 8804.FT.TO.CH       155   1.0000       5.0000      3.4441      1.0837
```

TABLE D-13

PRESCHOOL INTERPERSONAL PROBLEM SOLVING TEST

RESPONSE DISTRIBUTION FOR: ENGLISH-DOMINANT CHILDREN

VARIABLE	N	MINIMUM	MAXIMUM	MEAN	STD DEV
SOLUTNS	439	0.	8.0000	2.8018	1.8093
REL.GL.%	439	0.	1.0000	.20046	.40080
REL.GOAL	88	1.0000	7.0000	1.6818	1.1798
SUB.GL.%	439	0.	1.0000	.18907	.39201
SUB.GOAL	83	1.0000	8.0000	1.6627	1.3185
IRREL%	439	0.	1.0000	.50797	.50051
IRREL	223	1.0000	12.000	2.6188	1.9273
DKRNR%	439	0.	1.0000	.73804	.44020
DKRNR	324	1.0000	11.000	3.1667	1.9218
NPHYSOL	439	0.	6.0000	2.0615	1.4566
ASK%	439	0.	1.0000	.46469	.49932
PLEASE%	439	0.	1.0000	.42825	.49539
LOAN%	439	0.	1.0000	.59226 -1	.23632
FAIR%	439	0.	1.0000	.28246	.45071
TRADE%	439	0.	1.0000	.13667	.34390
AUTHINT%	439	0.	1.0000	.26196	.44020
TRICK%	439	0.	1.0000	.29613 -1	.16971
FINAGLE%	439	0.	1.0000	.11390 -1	.10623
MANAFF%	439	0.	1.0000	.75171 -1	.26397
MAD %	439	0.	1.0000	.18223 -1	.13391
WAIT%	439	0.	1.0000	.18679	.39019
FUTURE%	439	0.	1.0000	.10706	.30954
PHYSOL	439	0.	4.0000	.74032	.86589
FORCE%	439	0.	1.0000	.41002	.49240
PHYS%	439	0.	1.0000	.21185	.40908
DAMAGE %	439	0.	1.0000	.11390 -1	.10623
COMMAND%	439	0.	1.0000	.10706	.30954

TABLE D-14

BILINGUAL SYNTAX MEASURE-SPANISH (BSM-S)

RESPONSE DISTRIBUTION FOR: **SPANISH-DOMINANT CHILDREN**

N= 70

ITEM NUMBER	PERCENT OF CORRECT RESPONSE	PERCENT NO RESPONSE
1	91.4	4.3
2	90.0	7.1
3 (n= 70)	61.4	5.7
4 (n= 69)	51.4	1.4
5 (n= 69)	77.1	0.0
6 (n= 68)	41.4	5.7
7 (n= 68)	84.3	1.4
8 (n= 68)	91.4	0.0
11 (n= 68)	90.0	1.4
13 (n= 68)	82.9	1.4
14 (n= 68)	87.1	0.0
16 (n= 68)	12.9	5.7
19 (n= 68)	84.3	7.1
20 (n= 68)	74.3	5.7
21 (n= 68)	68.6	4.3
[illegible] (n= 67)	78.6	0.0
[illegible] (n= 67)	90.0	0.0
28 (n= 67)	10.0	2.9

TABLE D-15

BILINGUAL SYNTAX MEASURE-SPANISH (BSM-S)

RESPONSE DISTRIBUTION FOR: **ENGLISH-DOMINANT CHILDREN**

N= 8

ITEM NUMBER	PERCENT OF CORRECT RESPONSE	PERCENT NO RESPONSE
1	75.0	0.0
2	87.5	0.0
3 (n= 8)	37.5	12.5
4 (n= 8)	25.0	25.0
5 (n= 8)	62.5	12.5
6 (n= 7)	37.5	12.5
7 (n= 7)	87.5	0.0
8 (n= 7)	87.5	0.0
11 (n= 7)	62.5	12.5
13 (n= 7)	75.0	0.0
14 (n= 7)	62.5	0.0
16 (n= 7)	12.5	0.0
18 (n= 7)	37.5	12.5
20 (n= 7)	12.5	12.5
21 (n= 7)	75.0	0.0
22 (n= 7)	25.0	0.0
23 (n= 7)	62.5	0.0
24 (n= 7)	25.0	0.0

TABLE D-16

CONCEPTUAL GROUPING (ENGLISH VERSION)

RESPONSE DISTRIBUTION FOR: SPANISH-DOMINANT CHILDREN

N = 70

	PERCENT OF CHILDREN SCORING:			
ITEM NUMBER	TWO POINTS	ONE POINT	NO POINTS	NO RESPONSE
1	0.0	97.1	0.0	0.0
2	0.0	67.1	30.0	0.0
3	0.0	91.4	5.7	0.0
4	74.3	0.0	22.9	0.0
5 (n= 70)	17.1	4.3	75.7	0.0
6 (n= 70)	0.0	28.6	68.6	0.0
7 (n= 70)	0.0	41.4	55.7	0.0
8 (n= 58)	0.0	32.9	47.1	0.0
9 (n= 45)	11.4	32.9	15.7	1.4

TABLE D-17

SAY AND TELL (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: **SPANISH-DOMINANT CHILDREN**

N= 68

	PERCENT GIVING CORRECT RESPONSE	PERCENT NO RESPONSE
I. PENCIL		
NAME	89.7	1.5
COLOR	85.3	7.4
SHAPE	17.6	25.0
MATERIAL	47.1	22.1
FUNCTION	95.6	2.9
ANOTHER OBJECT FOR FUNCTION	5.9	11.8
ADDITIONAL INFORMATION	50.0	26.5
II. PENNIES		
LABEL	76.5	2.9
CLASS	23.5	2.9
COLOR	8.8	2.9
SHAPE	2.9	2.9
MATERIAL	0.0	2.9
FUNCTION	85.3	2.9
NUMBER	10.3	2.9
OTHER PHYSICAL CHARACTARISTICS	10.3	2.9
COMPARATIVE CHARACTERISTICS	0.0	2.9

TABLE D-18

VERBAL FLUENCY (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: **SPANISH-DOMINANT CHILDREN**

N = 68

	NO. OF EXAMPLES		
	MEAN	STANDARD DEVIATION	PERCENT NO RESPONSE
I. ANIMALS	1.6	1.9	19.1
II. THINGS TO EAT	2.0	1.3	5.9
III. NAMES	1.7	1.3	14.7
IV. TOYS	1.7	1.0	7.4

TABLE-D-19

ARM COORDINATION (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: **SPANISH-DOMINANT CHILDREN**

N = 67

	SCORE	PERCENT RESPONSE	PERCENT NO RESPONSE
I. BALL BOUNCING			2.99
	7	0.0	
	6	0.0	
	5	1.49	
	4	5.97	
	3	40.30	
	2	37.31	
	1	11.94	
	0	0.0	
II. BEANBAG CATCH GAME			
A. TOTAL TIMES BAG CAUGHT			0.0
	3	14.93	
	2	31.34	
	1	17.91	
	0	35.82	
B. BAG CAUGHT PREFERRED HAND			0.0
	3	2.99	
	2	8.96	
	1	16.42	
	0	71.64	
C. BAG CAUGHT NON-PREFERRED HAND			0.0
	3	4.48	
	2	5.97	
	1	8.96	
	0	80.60	
III. BEANBAG GAME			
A. PREFERRED HAND			1.49
	6	0.0	
	5	0.0	
	4	5.97	
	3	22.39	
	2	20.90	
	1	26.87	
	0	22.39	
B. NON-PREFERRED HAND			0.0
	6	0.0	
	5	0.0	
	4	1.49	
	3	7.46	
	2	25.37	
	1	22.39	
	0	19.40	

TABLE D-20

DRAW-A-CHILD (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: SPANISH-DOMINANT CHILDREN

N = 67

BODY PART	PERCENT OF CHILDREN HAVING BODY PART PRESENT IN DRAWING
HEAD	98.5
HAIR	49.3
EYES	92.5
NOSE	70.1
MOUTH	76.1
NECK	4.5
TRUNK	47.8
ARMS AND HANDS	55.2
ATTACHMENT OF ARMS	31.3
LEGS AND FEET	86.6
NO DRAWING	0.0

TABLE D-21

LEG COORDINATION(ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: SPANISH-DOMINANT CHILDREN

N = 67

	PERCENT RESPONSE	PERCENT NO RESPONSE
1.WALKING BACKWARD		3.0
5 OR MORE STEPS(2 PTS)	89.6	
2-4 STEPS(1 PT)	1.5	
0 PTS	6.0	
2.WALKING ON TIPTOE		0.0
5 OR MORE STEPS(2 PTS)	1.5	
2-4 STEPS(1 PT)	1.5	
0 PTS	97.0	
3.WALKING STRAIGHT LINE		0.0
2PTS	97.0	
1 PT	3.0	
0 PTS	0.0	
4.STANDING ON ONE FOOT		0.0
10 SEC OR MORE(2 PTS)	55.2	
3-9 SEC(1 PT)	41.8	
0-2 SEC(0 PTS)	3.0	
5.STANDING ON OTHER FOOT		0.0
10 SEC OR MORE(2 PTS)	43.3	
3-9 SEC(1 PT)	53.7	
0-2 SEC(0 PTS)	1.5	
6.SKIPPING		13.4
3 PTS	64.2	
2 PTS	10.4	
1 PT	1.5	
0 PTS	10.4	

TABLE D-22

PUPIL OBSERVATION CHECKLIST (POCL, FORM DC)

RESPONSE DISTRIBUTION FOR: SPANISH-DOMINANT CHILDREN

N = 70

CHILD'S TESTING BEHAVIOR	PERCENT OF CHILDREN IN RATINGS CATEGORIES 7	6	5	4	3	2	1	PERCENT NOT RATED
COOPERATIVE	1.4	14.3	24.3	44.3	7.1	5.7	0.0	2.9
SOCIABLE	4.3	7.1	18.6	32.9	18.6	14.3	1.4	2.9
INVOLVED	4.3	12.9	21.4	45.7	8.6	4.3	0.0	2.9
TALKATIVE	2.9	10.0	21.4	30.0	15.7	15.7	1.4	2.9
ATTENTIVE	1.4	21.4	15.7	41.4	5.7	10.0	0.0	4.3
ACTIVE	2.9	18.6	15.7	47.1	7.1	5.7	0.0	2.9
RELAXED	7.1	20.0	17.1	44.3	7.1	1.4	0.0	2.9
QUICK TO RESPOND	1.4	15.7	20.0	35.7	8.6	15.7	0.0	2.9
ATTEMPTS DIFFICULT TASKS	0.0	14.3	14.3	50.0	2.9	14.3	1.4	2.9
KEEPS TRYING	2.9	18.6	12.9	38.6	8.6	14.3	1.4	2.9
SELF-CONFIDENT	1.4	15.7	10.0	38.6	11.4	8.6	1.4	12.9
HANDICAP AFFECTED PERFORMANCE	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0

TABLE D-23

FACES

RESPONSE DISTRIBUTION FOR: SPANISH-DOMINANT CHILDREN

Variables are defined as follows:

801. Feelings about eating candy
802. Feelings about falling down
8801. Feelings about school this year
8802. Feelings about school next year
8803. Feelings about teacher
8804. Teachers feelings about child

VARIABLE	N	MINIMUM	MAXIMUM	MEAN	STD DEV
801.CANDY	73	1.0000	5.0000	3.4247	1.3115
802.FALL	74	1.0000	5.0000	2.3378	1.2194
8801.FSCHNOW	26	2.0000	5.0000	3.9231	1.0554
8802.FSCHNXY	26	1.0000	5.0000	3.2308	1.3359
8803.FCH.TO.T	25	1.5000	5.0000	3.4200	1.0567
8804.FT.TO.CH	26	2.0000	5.0000	3.5865	.91302

TABLE D-24

PRESCHOOL INTERPERSONAL PROBLEM SOLVING TEST

RESPONSE DISTRIBUTION FOR: SPANISH-DOMINANT CHILDREN

VARIABLE	N	MINIMUM	MAXIMUM	MEAN	STD DEV
SOLUTNS	75	0.	6.0000	2.3200	1.6778
REL.GL.%	75	0.	1.0000	.16000	.36907
REL.GOAL	12	1.0000	3.0000	1.6667	.77850
SUB.GL.%	75	0.	1.0000	.20000	.40269
SUB.GOAL	15	1.0000	3.0000	1.2667	.59362
IRREL%	75	0.	1.0000	.49333	.50332
IRREL	37	1.0000	9.0000	2.1892	1.4689
DKRNR%	75	0.	1.0000	.80000	.40269
DKRNR	60	1.0000	7.0000	3.1667	1.7190
NPHYSOL	75	0.	5.0000	1.6933	1.3353
ASK%	75	0.	1.0000	.30667	.46421
PLEASE%	75	0.	1.0000	.53333 -1	.22621
LOAN%	75	0.	1.0000	.40000	.49320
FAIR%	75	0.	1.0000	.20000	.40269
TRADE%	75	0.	1.0000	.18667	.39227
AUTHINT%	75	0.	1.0000	.32000	.46962
TRICK%	75	0.	0.	0.	
FINAGLE%	75	0.	1.0000	.13333 -1	.11547
MANAFF%	75	0.	1.0000	.53333 -1	.22621
MAD %	75	0.	1.0000	.13333 -1	.11547
WAIT%	75	0.	1.0000	.10667	.31077
FUTURE%	75	0.	1.0000	.40000 -1	.19728
PHYSOL	75	0.	3.0000	.62667	.86639
FORCE%	75	0.	1.0000	.34667	.47911
PHYS%	75	0.	1.0000	.16000	.36907
DAMAGE %	75	0.	1.0000	.26667 -1	.16219
COMMAND%	75	0.	1.0000	.93333 -1	.29286

TABLE D-25

BILINGUAL SYNTAX MEASURE-ENGLISH (BSM-E)

RESPONSE DISTRIBUTION FOR: THIRD GRADE

N= 30

ITEM NUMBER	PERCENT OF CORRECT RESPONSE	PERCENT NO RESPONSE
1	100.0	0.0
3	86.7	0.0
6 (n= 30)	66.7	0.0
7 (n= 30)	53.3	0.0
8 (n= 30)	73.3	0.0
9 (n= 30)	86.7	0.0
11 (n= 30)	93.3	0.0
13 (n= 30)	16.7	0.0
14 (n= 30)	70.0	0.0
15 (n= 30)	86.7	0.0
16 (n= 30)	90.0	0.0
17 (n= 30)	80.0	3.3
18 (n= 30)	96.7	0.0
19 (n= 30)	100.0	0.0
20 (n= 30)	93.3	0.0
22 (n= 30)	96.7	0.0
23 (n= 30)	30.0	0.0
24 (n= 30)	90.0	0.0

TABLE D-26

CONCEPTUAL GROUPING (ENGLISH VERSION)

RESPONSE DISTRIBUTION FOR: KINDERGARTEN

N = 26

	PERCENT OF CHILDREN SCORING:			
ITEM NUMBER	TWO POINTS	ONE POINT	NO POINTS	NO RESPONSE
1	0.0	100.0	0.0	0.0
2	0.0	96.2	3.8	0.0
3	0.0	96.2	3.8	0.0
4	80.8	11.5	7.7	0.0
5 (n= 26)	46.2	19.2	34.6	0.0
6 (n= 26)	0.0	34.6	65.4	0.0
7 (n= 26)	0.0	53.8	46.2	0.0
8 (n= 24)	0.0	53.8	38.5	0.0
9 (n= 22)	19.2	30.8	34.6	0.0

TABLE D-27

CONCEPTUAL GROUPING (ENGLISH VERSION)

RESPONSE DISTRIBUTION FOR: FIRST GRADE

N = 27

ITEM NUMBER	PERCENT OF CHILDREN SCORING: TWO POINTS	ONE POINT	NO POINTS	NO RESPONSE
1	0.0	100.0	0.0	0.0
2	0.0	100.0	0.0	0.0
3	0.0	100.0	0.0	0.0
4	96.3	3.7	0.0	0.0
5 (n= 27)	74.1	7.4	18.5	0.0
6 (n= 27)	0.0	63.0	37.0	0.0
7 (n= 27)	0.0	85.2	14.8	0.0
8 (n= 27)	0.0	74.1	25.9	0.0
9 (n= 26)	14.8	59.3	22.2	0.0

TABLE D-28

CONCEPTUAL GROUPING (ENGLISH VERSION)

RESPONSE DISTRIBUTION FOR: SECOND GRADE

N = 27

	PERCENT OF CHILDREN SCORING:			
ITEM NUMBER	TWO POINTS	ONE POINT	NO POINTS	NO RESPONSE
1	0.0	100.0	0.0	0.0
2	0.0	100.0	0.0	0.0
3	0.0	100.0	0.0	0.0
4	100.0	0.0	0.0	0.0
5 (n= 27)	85.2	0.0	14.8	0.0
6 (n= 27)	0.0	85.2	14.8	0.0
7 (n= 27)	0.0	92.6	7.4	0.0
8 (n= 27)	0.0	85.2	14.8	0.0
9 (n= 27)	7.4	74.1	18.5	0.0

TABLE D-29

CONCEPTUAL GROUPING (ENGLISH VERSION)

RESPONSE DISTRIBUTION FOR: THIRD GRADE **GEORGIA**

N = 24

	PERCENT OF CHILDREN SCORING:			
ITEM NUMBER	TWO POINTS	ONE POINT	NO POINTS	NO RESPONSE
1	0.0	100.0	0.0	0.0
2	0.0	100.0	0.0	0.0
3	0.0	100.0	0.0	0.0
4	95.8	0.0	4.2	0.0
5 (n= 24)	95.8	0.0	4.2	0.0
6 (n= 24)	0.0	87.5	12.5	0.0
7 (n= 24)	0.0	83.3	16.7	0.0
8 (n= 24)	0.0	87.5	12.5	0.0
9 (n= 24)	29.2	58.3	8.3	4.2

TABLE D-30

CONCEPTUAL GROUPING (ENGLISH VERSION)

RESPONSE DISTRIBUTION FOR: THIRD GRADE **MARYLAND**

N = 30

	PERCENT OF CHILDREN SCORING:			
ITEM NUMBER	TWO POINTS	ONE POINT	NO POINTS	NO RESPONSE
1	0.0	100.0	0.0	0.0
2	0.0	100.0	0.0	0.0
3	0.0	100.0	0.0	0.0
4	96.7	3.3	0.0	0.0
5 (n= 30)	86.7	0.0	13.3	0.0
6 (n= 30)	0.0	80.0	20.0	0.0
7 (n= 30)	0.0	86.7	13.3	0.0
8 (n= 30)	0.0	90.0	10.0	0.0
9 (n= 30)	43.3	40.0	16.7	0.0

TABLE **D-31**

SAY AND TELL (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: THIRD GRADE **MARYLAND**

N= 30

	PERCENT GIVING CORRECT RESPONSE	PERCENT NO RESPONSE
I. PENCIL		
NAME	100.0	0.0
COLOR	100.0	0.0
SHAPE	66.7	6.7
MATERIAL	86.7	0.0
FUNCTION	100.0	0.0
ANOTHER OBJECT FOR FUNCTION	56.7	0.0
ADDITIONAL INFORMATION	100.0	0.0
II. PENNIES		
LABEL	96.7	0.0
CLASS	13.3	0.0
COLOR	26.7	0.0
SHAPE	36.7	0.0
MATERIAL	13.3	0.0
FUNCTION	40.0	0.0
NUMBER	40.0	0.0
OTHER PHYSICAL CHARACTARISTICS	90.0	0.0
COMPARATIVE CHARACTERISTICS	3.3	0.0

TABLE D-32

VERBAL FLUENCY (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: KINDERGARTEN

N = 26

	NO. OF EXAMPLES		
	MEAN	STANDARD DEVIATION	PERCENT NO RESPONSE
I. ANIMALS	3.7	2.3	0.0
II. THINGS TO EAT	4.0	1.4	0.0
III. NAMES	3.6	1.2	0.0
IV. TOYS	2.8	1.3	0.0

TABLE D-33

VERBAL FLUENCY (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: FIRST GRADE

N = 27

	NO. OF EXAMPLES		
	MEAN	STANDARD DEVIATION	PERCENT NO RESPONSE
I. ANIMALS	4.5	2.1	7.4
II. THINGS TO EAT	4.4	1.6	0.0
III. NAMES	4.4	1.5	0.0
IV. TOYS	3.2	1.3	0.0

TABLE D-34

VERBAL FLUENCY (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: SECOND GRADE

N = 27

	NO. OF EXAMPLES		
	MEAN	STANDARD DEVIATION	PERCENT NO RESPONSE
I. ANIMALS	6.4	1.8	0.0
II. THINGS TO EAT	5.1	1.5	0.0
III. NAMES	5.4	1.8	0.0
IV. TOYS	3.8	1.8	0.0

TABLE D-35

VERBAL FLUENCY (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: THIRD GRADE **GEORGIA**

N = 24

	NO. OF EXAMPLES		
	MEAN	STANDARD DEVIATION	PERCENT NO RESPONSE
I. ANIMALS	6.8	1.7	0.0
II. THINGS TO EAT	5.7	1.6	0.0
III. NAMES	5.3	2.0	0.0
IV. TOYS	4.0	1.0	0.0

TABLE D-36

VERBAL FLUENCY (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: THIRD GRADE **MARYLAND**

N = 30

	NO. OF EXAMPLES MEAN	STANDARD DEVIATION	PERCENT NO RESPONSE
I. ANIMALS	6.5	2.3	0.0
II. THINGS TO EAT	6.2	2.1	0.0
III. NAMES	5.7	2.4	0.0
IV. TOYS	4.1	1.5	0.0

TABLE D-37

VERBAL MEMORY (MSCA) (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: KINDERGARTEN

N = 24

I. ITEM	NUMBER OF WORDS REPEATED 4	3	2	1	0	CORRECT SEQUENCE
1.TOY-CHAIR-LIGHT	0	83	0	4	8	88 %
2.DOLL-DARK-COAT	0	79	17	0	0	92 %
3.AFTER-COLOR-FUNNY -TODAY	88	8	0	0	0	75 %
4.AROUND-BECAUSE-UNDER -NEVER	46	42	4	4	0	83 %

	MC CARTHY SCORE 4	3	2	1	0	NO RESPONSE
1.TOY-CHAIR-LIGHT	0	83	0	4	8	4 %
2.DOLL-DARK-COAT	0	75	21	0	0	4 %
3.AFTER-COLOR-FUNNY -TODAY	67	29	0	0	0	4 %
4.AROUND-BECAUSE-UNDER -NEVER	46	29	17	4	0	4 %

III. STORY PART	CREDIT FOR PARTS OF THE STORY ONE POINT	TWO POINTS	NO RESPONSE
1.TERM USED FOR BOB	91.7%	4.2%	4.2%
2.TERM USED FOR WOMAN	91.7	4.2	4.2
3.TERM USED FOR LETTERS	91.7	4.2	4.2
4.BOB WALKING TO STORE	37.5	58.3	4.2
5.BOB SAW WOMAN	50.0	45.8	4.2
6.WIND BLEW LETTERS	79.2	16.7	4.2
7.BOB SHOUTED, "..."	41.7	54.2	4.2
8.BOB WAS CAREFUL	62.5	33.3	4.2
9.BOB PICKED UP LETTERS	83.3	12.5	4.2
10.WOMAN WAS HAPPY	20.8	75.0	4.2
11.WOMAN THANKED BOB	33.3	62.5	4.2

TABLE D-38

VERBAL MEMORY (MSCA) (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: FIRST GRADE

N = 27

I. ITEM	NUMBER OF WORDS REPEATED 4	3	2	1	0	CORRECT SEQUENCE
1.TOY-CHAIR-LIGHT	0	96	0	4	0	100 %
2.DOLL-DARK-COAT	0	85	11	0	4	96 %
3.AFTER-COLOR-FUNNY -TODAY	81	19	0	0	0	89 %
4.AROUND-BECAUSE-UNDER -NEVER	59	22	15	4	0	85 %

	MC CARTHY SCORE 4	3	2	1	0	NO RESPONSE
1.TOY-CHAIR-LIGHT	0	96	0	4	0	0 %
2.DOLL-DARK-COAT	0	85	11	0	4	0 %
3.AFTER-COLOR-FUNNY -TODAY	74	22	4	0	0	0 %
4.AROUND-BECAUSE-UNDER -NEVER	52	26	15	7	0	0 %

III. STORY PART	CREDIT FOR PARTS OF THE STORY ONE POINT	TWO POINTS	NO RESPONSE
1.TERM USED FOR BOB	88.9%	3.7%	7.4%
2.TERM USED FOR WOMAN	88.9	3.7	7.4
3.TERM USED FOR LETTERS	88.9	3.7	7.4
4.BOB WALKING TO STORE	33.3	59.3	7.4
5.BOB SAW WOMAN	37.0	55.6	7.4
6.WIND BLEW LETTERS	74.1	18.5	7.4
7.BOB SHOUTED. "..."	29.6	63.0	7.4
8.BOB WAS CAREFUL	40.7	51.9	7.4
9.BOB PICKED UP LETTERS	81.5	11.1	7.4
10.WOMAN WAS HAPPY	33.3	59.3	7.4
11.WOMAN THANKED BOB	33.3	59.3	7.4

TABLE D-39

VERBAL MEMORY (MSCA) (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: SECOND GRADE

N = 27

I. ITEM	NUMBER OF WORDS REPEATED 4	3	2	1	0	CORRECT SEQUENCE
1.TOY-CHAIR-LIGHT	0	100	0	0	0	100 %
2.DOLL-DARK-COAT	0	96	4	0	0	100 %
3.AFTER-COLOR-FUNNY -TODAY	96	4	0	0	0	74 %
4.AROUND-BECAUSE-UNDER -NEVER	67	22	7	4	0	81 %

	MC CARTHY SCORE 4	3	2	1	0	NO RESPONSE
1.TOY-CHAIR-LIGHT	0	100	0	0	0	0 %
2.DOLL-DARK-COAT	0	96	4	0	0	0 %
3.AFTER-COLOR-FUNNY -TODAY	70	30	0	0	0	0 %
4.AROUND-BECAUSE-UNDER NEVER	59	19	19	4	0	0 %

III. STORY PART	CREDIT FOR PARTS OF THE STORY ONE POINT	TWO POINTS	NO RESPONSE
1.TERM USED FOR BOB	96.3%	0.0%	3.7%
2.TERM USED FOR WOMAN	96.3	0.0	3.7
3.TERM USED FOR LETTERS	92.6	3.7	3.7
4.BOB WALKING TO STORE	22.2	74.1	3.7
5.BOB SAW WOMAN	66.7	29.6	3.7
6.WIND BLEW LETTERS	81.5	14.8	3.7
7.BOB SHOUTED, "..."	59.3	37.0	3.7
8.BOB WAS CAREFUL	66.7	29.6	3.7
9.BOB PICKED UP LETTERS	77.8	18.5	3.7
10.WOMAN WAS HAPPY	14.8	81.5	3.7
11.WOMAN THANKED BOB	55.6	40.7	3.7

TABLE D-40

VERBAL MEMORY (MSCA) (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: THIRD GRADE **GEORGIA**

N = 24

I. ITEM	NUMBER OF WORDS REPEATED 4	3	2	1	0	CORRECT SEQUENCE
1.TOY-CHAIR-LIGHT	0	100	0	0	0	100 %
2.DOLL-DARK-COAT	0	100	0	0	0	100 %
3.AFTER-COLOR-FUNNY -TODAY	100	0	0	0	0	96 %
4.AROUND-BECAUSE-UNDER -NEVER	79	13	4	4	0	79 %

	MC CARTHY SCORE 4	3	2	1	0	NO RESPONSE
1.TOY-CHAIR-LIGHT	0	100	0	0	0	0 %
2.DOLL-DARK-COAT	0	100	0	0	0	0 %
3.AFTER-COLOR-FUNNY -TODAY	96	4	0	0	0	0 %
4.AROUND-BECAUSE-UNDER -NEVER	63	25	8	4	0	0 %

III. STORY PART	CREDIT FOR PARTS OF THE STORY ONE POINT	TWO POINTS	NO RESPONSE
1.TERM USED FOR BOB	100.0%	0.0%	0.0%
2.TERM USED FOR WOMAN	100.0	0.0	0.0
3.TERM USED FOR LETTERS	100.0	0.0	0.0
4.BOB WALKING TO STORE	37.5	62.5	0.0
5.BOB SAW WOMAN	66.7	33.3	0.0
6.WIND BLEW LETTERS	95.8	4.2	0.0
7.BOB SHOUTED, "..."	54.2	45.8	0.0
8.BOB WAS CAREFUL	70.8	29.2	0.0
9.BOB PICKED UP LETTERS	83.3	16.7	0.0
10.WOMAN WAS HAPPY	37.5	62.5	0.0
11.WOMAN THANKED BOB	70.8	29.2	0.0

TABLE **D-41**

VERBAL MEMORY (MSCA) (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: THIRD GRADE **MARYLAND**

N = 30

I. ITEM	NUMBER OF WORDS REPEATED 4	3	2	1	0	CORRECT SEQUENCE
1.TOY-CHAIR-LIGHT	0	97	3	0	0	100 %
2.DOLL-DARK-COAT	0	93	7	0	0	100 %
3.AFTER-COLOR-FUNNY -TODAY	97	3	0	0	0	87 %
4.AROUND-BECAUSE-UNDER -NEVER	90	10	0	0	0	93 %

	MC CARTHY SCORE 4	3	2	1	0	NO RESPONSE
1.TOY-CHAIR-LIGHT	0	97	3	0	0	0 %
2.DOLL-DARK-COAT	0	93	7	0	0	0 %
3.AFTER-COLOR-FUNNY -TODAY	83	17	0	0	0	0 %
4.AROUND-BECAUSE-UNDER -NEVER	87	10	3	0	0	0 %

III. STORY PART	CREDIT FOR PARTS OF THE STORY ONE POINT	TWO POINTS	NO RESPONSE
1.TERM USED FOR BOB	100.0%	0.0%	0.0%
2.TERM USED FOR WOMAN	96.7	3.3	0.0
3.TERM USED FOR LETTERS	100.0	0.0	0.0
4.BOB WALKING TO STORE	40.0	60.0	0.0
5.BOB SAW WOMAN	56.7	43.3	0.0
6.WIND BLEW LETTERS	83.3	16.7	0.0
7.BOB SHOUTED, "..."	73.3	26.7	0.0
8.BOB WAS CAREFUL	66.7	33.3	0.0
9.BOB PICKED UP LETTERS	90.0	10.0	0.0
10.WOMAN WAS HAPPY	26.7	73.3	0.0
11.WOMAN THANKED BOB	63.3	36.7	0.0

TABLE D-42

ARM COORDINATION (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: KINDERGARTEN

N = 26

	PERCENT RESPONSE	PERCENT NO RESPONSE
I. BALL BOUNCING		
SCORE		0.0
7	3.85	
6	0.0	
5	3.85	
4	7.69	
3	34.62	
2	26.92	
1	23.08	
0	0.0	
II. BEANBAG CATCH GAME		
A. TOTAL TIMES BAG CAUGHT		0.0
3	30.77	
2	23.08	
1	30.77	
0	15.38	
B. BAG CAUGHT PREFERRED HAND		0.0
3	3.85	
2	15.38	
1	26.92	
0(N= 26)	53.85	
C. BAG CAUGHT NON PREFERRED HAND		0.0
3	3.85	
2	3.85	
1	19.23	
0(N= 26)	73.08	
III. BEANBAG GAME		
A. PREFERRED HAND		0.0
6	0.0	
5	0.0	
4	11.54	
3	23.08	
2	26.92	
1	23.08	
0	15.38	
B. NON-PREFERRED HAND		0.0
6	0.0	
5	0.0	
4	3.85	
3	19.23	
2	15.38	
1	23.08	
0	23.08	

TABLE D-43

ARM COORDINATION (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: FIRST GRADE

N = 27

	PERCENT RESPONSE	PERCENT NO RESPONSE
I. BALL BOUNCING		
SCORE		0.0
7	22.22	
6	3.70	
5	7.41	
4	7.41	
3	37.04	
2	18.52	
1	3.70	
0	0.0	
II. BEAN BAG CATCH GAME		
A. TOTAL TIMES BAG CAUGHT		0.0
3	18.52	
2	51.85	
1	22.22	
0	7.41	
B. BAG CAUGHT PREFERRED HAND		0.0
3	14.81	
2	33.33	
1	29.63	
0(N= 27)	22.22	
C. BAG CAUGHT NON-PREFERRED HAND		0.0
3	14.81	
2	25.93	
1	18.52	
0(N= 27)	40.74	
III BEANBAG GAME		
A. PREFERRED HAND		0.0
6	0.0	
5	3.70	
4	3.70	
3	7.41	
2	22.22	
1	37.04	
0	25.93	
B. NON-PREFERRED HAND		0.0
6	0.0	
5	3.70	
4	7.41	
3	11.11	
2	7.41	
1	29.63	
0	14.81	

TABLE D-44

ARM COORDINATION (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: SECOND GRADE

N = 27

	PERCENT RESPONSE	PERCENT NO RESPONSE
I. BALL BOUNCING		
SCORE		0.0
7	37.04	
6	11.11	
5	11.11	
4	11.11	
3	18.52	
2	0.0	
1	11.11	
0	0.0	
II. BEANBAG CATCH GAME		
A. TOTAL TIMES BAG CAUGHT		0.0
3	48.15	
2	22.22	
1	14.81	
0	14.81	
B. BAG CAUGHT PREFERRED HAND		0.0
3	25.93	
2	18.52	
1	25.93	
0(N= 27)	29.63	
C. BAG CAUGHT NON-PREFERRED HAND		0.0
3	25.93	
2	25.93	
1	14.81	
0(N= 27)	33.33	
III. BEANBAG GAME		
A. PREFERRED HAND		0.0
6	0.0	
5	3.70	
4	18.52	
3	33.33	
2	25.93	
1	11.11	
0	7.41	
B. NON-PREFERRED HAND		0.0
6	0.0	
5	3.70	
4	14.81	
3	22.22	
2	25.93	
1	25.93	
0	0.0	

TABLE D-45

ARM COORDINATION (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: THIRD GRADE **GEORGIA**

N = 24

	PERCENT RESPONSE	PERCENT NO RESPONSE
I. BALL BOUNCING		
SCORE		0.0
7	66.67	
6	4.17	
5	4.17	
4	16.67	
3	8.33	
2	0.0	
1	0.0	
0	0.0	
II. BEANBAG CATCH GAME		
A. TOTAL TIMES BAG CAUGHT		0.0
3	83.33	
2	16.67	
1	0.0	
0	0.0	
B. BAG CAUGHT PREFERRED HAND		0.0
3	41.67	
2	29.17	
1	20.83	
0(N= 24)	8.33	
C. BAG CAUGHT NON-PREFERRED HAND		0.0
3	50.00	
2	29.17	
1	12.50	
0(N= 24)	8.33	
III. BEANBAG GAME		
A. PREFERRED HAND		0.0
6	0.0	
5	8.33	
4	29.17	
3	33.33	
2	12.50	
1	16.67	
0	0.0	
B. NON-PREFERRED HAND		0.0
6	0.0	
5	8.33	
4	12.50	
3	20.83	
2	29.17	
1	25.00	
0	4.17	

TABLE D-46

ARM COORDINATION (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: THIRD GRADE **MARYLAND**

N = 30

	PERCENT RESPONSE	PERCENT NO RESPONSE
I. BALL BOUNCING		
SCORE		0.0
7	76.67	
6	0.0	
5	10.00	
4	0.0	
3	13.33	
2	0.0	
1	0.0	
0	0.0	
II. BEANBAG CATCH GAME		
A. TOTAL TIMES BAG CAUGHT		0.0
3	66.67	
2	23.33	
1	10.00	
0	0.0	
B. BAG CAUGHT PREFERRED HAND		0.0
3	50.00	
2	26.67	
1	16.67	
0(N= 30)	6.67	
C. BAG CAUGHT NON-PREFERRED HAND		0.0
3	43.33	
2	36.67	
1	10.00	
0(N= 30)	10.00	
III. BEANBAG GAME		
A. PREFERRED HAND		0.0
6	0.0	
5	13.33	
4	40.00	
3	16.67	
2	20.00	
1	3.33	
0	6.67	
B. NON-PREFERRED HAND		0.0
6	0.0	
5	0.0	
4	10.00	
3	33.33	
2	26.67	
1	20.00	
0	3.33	

TABLE D-47

DRAW-A-CHILD (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: KINDERGARTEN

N = 26

BODY PART	PERCENT OF CHILDREN HAVING BODY PART PRESENT IN DRAWING
HEAD	100.0
HAIR	76.9
EYES	100.0
NOSE	65.4
MOUTH	88.5
NECK	19.2
TRUNK	73.1
ARMS AND HANDS	88.5
ATTACHMENT OF ARMS	57.7
LEGS AND FEET	100.0
NO DRAWING	0.0

TABLE D-48

DRAW-A-CHILD (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: FIRST GRADE

N = 27

BODY PART	PERCENT OF CHILDREN HAVING BODY PART PRESENT IN DRAWING
HEAD	100.0
HAIR	92.6
EYES	100.0
NOSE	88.9
MOUTH	92.6
NECK	29.6
TRUNK	92.6
ARMS AND HANDS	92.6
ATTACHMENT OF ARMS	81.5
LEGS AND FEET	96.3
NO DRAWING	0.0

TABLE D-49

DRAW-A-CHILD (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: SECOND GRADE

N = 27

BODY PART	PERCENT OF CHILDREN HAVING BODY PART PRESENT IN DRAWING
HEAD	100.0
HAIR	85.2
EYES	100.0
NOSE	92.6
MOUTH	96.3
NECK	51.9
TRUNK	88.9
ARMS AND HANDS	100.0
ATTACHMENT OF ARMS	88.9
LEGS AND FEET	96.3
NO DRAWING	0.0

TABLE D-50

DRAW-A-CHILD (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: THIRD GRADE GEORGIA

N = 24

BODY PART	PERCENT OF CHILDREN HAVING BODY PART PRESENT IN DRAWING
HEAD	100.0
HAIR	95.8
EYES	100.0
NOSE	95.8
MOUTH	100.0
NECK	58.3
TRUNK	100.0
ARMS AND HANDS	95.8
ATTACHMENT OF ARMS	95.8
LEGS AND FEET	95.8
NO DRAWING	0.0

TABLE D-51

DRAW-A-CHILD (ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: THIRD GRADE **MARYLAND**

N = 29

BODY PART	PERCENT OF CHILDREN HAVING BODY PART PRESENT IN DRAWING
HEAD	100.0
HAIR	93.1
EYES	100.0
NOSE	100.0
MOUTH	100.0
NECK	62.1
TRUNK	100.0
ARMS AND HANDS	100.0
ATTACHMENT OF ARMS	96.6
LEGS AND FEET	100.0
NO DRAWING	0.0

TABLE D-52

LEG COORDINATION(ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: KINDERGARTEN

N = 24

	PERCENT RESPONSE	PERCENT NO RESPONSE
1.WALKING BACKWARD		0.0
5 OR MORE STEPS(2 PTS)	95.8	
2-4 STEPS(1 PT)	0.0	
0 PTS	4.2	
2.WALKING ON TIPTOE		0.0
5 OR MORE STEPS(2 PTS)	0.0	
2-4 STEPS(1 PT)	4.2	
0 PTS	95.8	
3.WALKING STRAIGHT LINE		0.0
2PTS	95.8	
1 PT	0.0	
0 PTS	4.2	
4.STANDING ON ONE FOOT		0.0
10 SEC OR MORE(2 PTS)	66.7	
3-9 SEC(1 PT)	25.0	
0-2 SEC(0 PTS)	8.3	
5.STANDING ON OTHER FOOT		0.0
10 SEC OR MORE(2 PTS)	70.8	
3-9 SEC(1 PT)	20.8	
0-2 SEC(0 PTS)	4.2	
6.SKIPPING		4.2
3 PTS	87.5	
2 PTS	4.2	
1 PT	0.0	
0 PTS	4.2	

TABLE D-53

LEG COORDINATION(ENGLISH OR SPANISH VERSION)

RESPONSE DISTRIBUTION FOR: THIRD GRADE **MARYLAND**

N = 30

	PERCENT RESPONSE	PERCENT NO RESPONSE
1.WALKING BACKWARD		0.0
5 OR MORE STEPS(2 PTS)	100.0	
2-4 STEPS(1 PT)	0.0	
0 PTS	0.0	
2.WALKING ON TIPTOE		0.0
5 OR MORE STEPS(2 PTS)	0.0	
2-4 STEPS(1 PT)	0.0	
0 PTS	100.0	
3.WALKING STRAIGHT LINE		0.0
2PTS	100.0	
1 PT	0.0	
0 PTS	0.0	
4.STANDING ON ONE FOOT		0.0
10 SEC OR MORE(2 PTS)	93.3	
3-9 SEC(1 PT)	6.7	
0-2 SEC(0 PTS)	0.0	
5.STANDING ON OTHER FOOT		0.0
10 SEC OR MORE(2 PTS)	100.0	
3-9 SEC(1 PT)	0.0	
0-2 SEC(0 PTS)	0.0	
6.SKIPPING		0.0
3 PTS	100.0	
2 PTS	0.0	
1 PT	0.0	
0 PTS	0.0	

TABLE D-54

PUPIL OBSERVATION CHECKLIST (POCL, FORM DC)

RESPONSE DISTRIBUTION FOR: KINDERGARTEN

N = 26

	PERCENT OF CHILDREN IN RATINGS CATEGORIES							PERCENT NOT RATED
	7	6	5	4	3	2	1	
CHILD'S TESTING BEHAVIOR								
COOPERATIVE	0.0	0.0	7.7	76.9	0.0	0.0	0.0	15.4
SOCIABLE	0.0	0.0	15.4	65.4	3.8	0.0	0.0	15.4
INVOLVED	0.0	0.0	0.0	84.6	0.0	0.0	0.0	15.4
TALKATIVE	0.0	3.8	7.7	69.2	3.8	0.0	0.0	15.4
ATTENTIVE	0.0	0.0	0.0	80.8	3.8	0.0	0.0	15.4
ACTIVE	0.0	7.7	0.0	76.9	0.0	0.0	0.0	15.4
RELAXED	0.0	0.0	0.0	84.6	0.0	0.0	0.0	15.4
QUICK TO RESPOND	0.0	0.0	3.8	73.1	7.7	0.0	0.0	15.4
ATTEMPTS DIFFICULT TASKS	0.0	0.0	0.0	80.8	3.8	0.0	0.0	15.4
KEEPS TRYING	0.0	0.0	0.0	84.6	0.0	0.0	0.0	15.4
SELF-CONFIDENT	0.0	0.0	0.0	76.9	0.0	0.0	0.0	23.1
HANDICAP AFFECTED PERFORMANCE	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0

TABLE D-55

PUPIL OBSERVATION CHECKLIST (POCL, FORM DC)

RESPONSE DISTRIBUTION FOR: FIRST GRADE

N = 27

CHILD'S TESTING BEHAVIOR	PERCENT OF CHILDREN IN RATINGS CATEGORIES 7	6	5	4	3	2	1	PERCENT NOT RATED
COOPERATIVE	0.0	0.0	22.2	51.9	14.8	0.0	0.0	11.1
SOCIABLE	0.0	0.0	18.5	51.9	11.1	7.4	0.0	11.1
INVOLVED	0.0	0.0	0.0	85.2	3.7	0.0	0.0	11.1
TALKATIVE	0.0	0.0	14.8	51.9	14.8	3.7	3.7	11.1
ATTENTIVE	0.0	0.0	0.0	77.8	11.1	0.0	0.0	11.1
ACTIVE	0.0	0.0	3.7	70.4	14.8	0.0	0.0	11.1
RELAXED	0.0	0.0	3.7	85.2	0.0	0.0	0.0	11.1
QUICK TO RESPOND	0.0	0.0	0.0	77.8	3.7	7.4	0.0	11.1
ATTEMPTS DIFFICULT TASKS	0.0	0.0	0.0	81.5	7.4	0.0	0.0	11.1
KEEPS TRYING	0.0	0.0	0.0	77.8	11.1	0.0	0.0	11.1
SELF-CONFIDENT	0.0	0.0	0.0	77.8	11.1	0.0	0.0	11.1
HANDICAP AFFECTED PERFORMANCE	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0

TABLE D-56

PUPIL OBSERVATION CHECKLIST (POCL, FORM DC)

RESPONSE DISTRIBUTION FOR: SECOND GRADE

N = 27

	PERCENT OF CHILDREN IN RATINGS CATEGORIES							PERCENT NOT RATED
	7	6	5	4	3	2	1	
CHILD'S TESTING BEHAVIOR								
COOPERATIVE	3.7	3.7	44.4	44.4	3.7	0.0	0.0	0.0
SOCIABLE	0.0	3.7	48.1	37.0	7.4	0.0	3.7	0.0
INVOLVED	0.0	0.0	29.6	63.0	7.4	0.0	0.0	0.0
TALKATIVE	0.0	7.4	18.5	48.1	22.2	0.0	3.7	0.0
ATTENTIVE	0.0	7.4	18.5	63.0	11.1	0.0	0.0	0.0
ACTIVE	0.0	0.0	7.4	85.2	7.4	0.0	0.0	0.0
RELAXED	0.0	0.0	14.8	77.8	3.7	0.0	0.0	3.7
QUICK TO RESPOND	0.0	7.4	11.1	66.7	11.1	0.0	0.0	3.7
ATTEMPTS DIFFICULT TASKS	0.0	11.1	3.7	81.5	3.7	0.0	0.0	0.0
KEEPS TRYING	0.0	7.4	7.4	77.8	7.4	0.0	0.0	0.0
SELF-CONFIDENT	0.0	3.7	11.1	74.1	11.1	0.0	0.0	0.0
HANDICAP AFFECTED PERFORMANCE	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0

TABLE D-57

PUPIL OBSERVATION CHECKLIST (POCL, FORM DC)

RESPONSE DISTRIBUTION FOR: THIRD GRADE GEORGIA

N = 24

	PERCENT OF CHILDREN IN RATINGS CATEGORIES							PERCENT NOT RATED
	7	6	5	4	3	2	1	
CHILD'S TESTING BEHAVIOR								
COOPERATIVE	0.0	4.2	12.5	75.0	0.0	0.0	0.0	8.3
SOCIABLE	0.0	4.2	4.2	75.0	4.2	4.2	0.0	8.3
INVOLVED	0.0	4.2	4.2	83.3	0.0	0.0	0.0	8.3
TALKATIVE	0.0	0.0	4.2	75.0	8.3	4.2	0.0	8.3
ATTENTIVE	0.0	0.0	8.3	83.3	0.0	0.0	0.0	8.3
ACTIVE	0.0	0.0	0.0	91.7	0.0	0.0	0.0	8.3
RELAXED	0.0	0.0	4.2	87.5	0.0	0.0	0.0	8.3
QUICK TO RESPOND	0.0	0.0	4.2	87.5	0.0	0.0	0.0	8.3
ATTEMPTS DIFFICULT TASKS	0.0	0.0	4.2	87.5	0.0	0.0	0.0	8.3
KEEPS TRYING	0.0	0.0	4.2	87.5	0.0	0.0	0.0	8.3
SELF-CONFIDENT	0.0	0.0	4.2	83.3	0.0	0.0	0.0	12.5
HANDICAP AFFECTED PERFORMANCE	0.0	0.0	0.0	0.0	0.0	0.0	0.0	100.0

TABLE D-58

PDC CHILD RATING SCALE (ENGLISH ONLY RATED BY TEACHER)
RESPONSE DISTRIBUTION FOR: KINDERGARTEN
TOTAL NUMBER OF RATINGS: 13

LENGTH OF TIME TEACHER HAD KNOWN CHILD:	PERCENT RESPONDING
LESS THAN MONTH	0.0
ONE-SIX MONTHS	30.8
SIX-TWELVE MONTHS	61.5
MORE THAN 12 MONTHS	0.0
HOW WELL TEACHER KNOWS CHILD	
NOT WELL	0.0
AVERAGE	84.6
BETTER THAN AVERAGE	7.7

I. SOCIAL COMPETENCE

	FREQUENTLY	MORE THAN AVERAGE	AVERAGE	LESS THAN AVERAGE	RARELY	NO RESPONSE
1.	0.0	15.4	53.8	23.1	0.0	0.0
2.	0.0	0.0	38.5	53.8	0.0	0.0
3.	0.0	15.4	61.5	15.4	0.0	0.0
4.	0.0	15.4	69.2	7.7	0.0	0.0
5.	0.0	0.0	92.3	0.0	0.0	0.0
6.	0.0	0.0	76.9	15.4	0.0	0.0
7.	0.0	0.0	92.3	0.0	0.0	0.0
8.	0.0	30.8	46.2	15.4	0.0	0.0
9.	0.0	30.8	61.5	0.0	0.0	0.0
10.	0.0	7.7	69.2	15.4	0.0	0.0
11.	0.0	0.0	38.5	53.8	0.0	0.0
12.	0.0	0.0	84.6	7.7	0.0	0.0
13.	0.0	0.0	38.5	53.8	0.0	0.0
14.	0.0	15.4	53.8	15.4	7.7	0.0
15.	0.0	0.0	84.6	0.0	0.0	0.0
16.	0.0	30.8	46.2	15.4	0.0	0.0
17.	0.0	30.8	61.5	0.0	0.0	0.0
18.	0.0	30.8	61.5	0.0	0.0	0.0
19.	0.0	7.7	61.5	15.4	7.7	0.0
20.	0.0	7.7	84.6	0.0	0.0	0.0
21.	0.0	23.1	61.5	7.7	0.0	0.0
22.	0.0	38.5	46.2	7.7	0.0	0.0
23.	0.0	15.4	61.5	15.4	0.0	0.0
24.	0.0	15.4	46.2	23.1	7.7	0.0

II. LEARNING HOW TO LEARN

	FREQUENTLY	MORE THAN AVERAGE	AVERAGE	LESS THAN AVERAGE	RARELY	NO RESPONSE
25.	0.0	23.1	53.8	7.7	0.0	0.0
26.	0.0	7.7	61.5	7.7	7.7	0.0
27.	0.0	15.4	61.5	7.7	0.0	0.0
28.	0.0	0.0	53.8	23.1	7.7	0.0
29.	0.0	0.0	53.8	23.1	7.7	0.0
30.	7.7	7.7	53.8	15.4	0.0	0.0
31.	7.7	15.4	46.2	15.4	0.0	0.0
32.	7.7	15.4	46.2	7.7	7.7	0.0
33.	0.0	23.1	53.8	7.7	0.0	0.0
34.	0.0	23.1	46.2	7.7	7.7	0.0
35.	0.0	7.7	69.2	7.7	0.0	0.0
36.	0.0	0.0	69.2	7.7	7.7	0.0
37.	0.0	7.7	61.5	7.7	7.7	0.0
38.	0.0	23.1	46.2	7.7	7.7	0.0
39.	100.0	0.0	0.0	0.0	0.0	0.0

TABLE D-59

PDC CHILD RATING SCALE (ENGLISH RATED ONLY BY TEACHER)
RESPONSE DISTRIBUTION FOR: FIRST GRADE
TOTAL NUMBER OF RATINGS: 14

LENGTH OF TIME TEACHER HAD KNOWN CHILD:	PERCENT RESPONDING
LESS THAN MONTH	0.0
ONE-SIX MONTHS	0.0
SIX-TWELVE MONTHS	92.9
MORE THAN 12 MONTHS	0.0
HOW WELL TEACHER KNOWS CHILD	
NOT WELL	0.0
AVERAGE	78.6
BETTER THAN AVERAGE	14.3

I. SOCIAL COMPETENCE

	FREQUENTLY	MORE THAN AVERAGE	AVERAGE	LESS THAN AVERAGE	RARELY	NO RESPONSE
1.	7.1	14.3	42.9	14.3	14.3	0.0
2.	7.1	14.3	50.0	0.0	21.4	0.0
3.	0.0	14.3	50.0	14.3	14.3	0.0
4.	0.0	28.6	50.0	7.1	7.1	0.0
5.	0.0	28.6	42.9	7.1	14.3	0.0
6.	7.1	21.4	57.1	0.0	7.1	0.0
7.	7.1	21.4	57.1	7.1	0.0	0.0
8.	21.4	7.1	42.9	14.3	7.1	0.0
9.	21.4	7.1	57.1	7.1	0.0	0.0
10.	0.0	7.1	64.3	14.3	7.1	0.0
11.	7.1	7.1	42.9	14.3	21.4	0.0
12.	9.0	0.0	78.6	7.1	7.1	0.0
13.	7.1	7.1	35.7	14.3	28.6	0.0
14.	0.0	7.1	64.3	21.4	0.0	0.0
15.	0.0	14.3	78.6	0.0	0.0	0.0
16.	14.3	7.1	64.3	0.0	7.1	0.0
17.	0.0	0.0	85.7	7.1	0.0	0.0
18.	0.0	0.0	92.9	0.0	0.0	0.0
19.	7.1	7.1	50.0	28.6	0.0	0.0
20.	0.0	0.0	71.4	7.1	7.1	0.0
21.	14.3	7.1	50.0	21.4	0.0	0.0
22.	14.3	14.3	57.1	0.0	7.1	0.0
23.	7.1	21.4	42.9	14.3	7.1	0.0
24.	7.1	21.4	35.7	21.4	7.1	0.0

II. LEARNING HOW TO LEARN

	FREQUENTLY	MORE THAN AVERAGE	AVERAGE	LESS THAN AVERAGE	RARELY	NO RESPONSE
25.	14.3	14.3	42.9	14.3	7.1	0.0
26.	7.1	21.4	21.4	21.4	7.1	0.0
27.	0.0	21.4	42.9	14.3	14.3	0.0
28.	0.0	14.3	64.3	14.3	0.0	0.0
29.	0.0	21.4	14.3	14.3	35.7	0.0
30.	21.4	0.0	50.0	14.3	0.0	0.0
31.	7.1	21.4	28.6	14.3	21.4	0.0
32.	14.3	7.1	35.7	35.7	0.0	0.0
33.	21.4	14.3	42.9	7.1	0.0	0.0
34.	21.4	14.3	57.1	0.0	0.0	0.0
35.	0.0	14.3	42.9	14.3	21.4	0.0
36.	0.0	7.1	57.1	21.4	7.1	0.0
37.	7.1	21.4	42.9	21.4	0.0	0.0
38.	7.1	14.3	42.9	21.4	7.1	0.0
39.	100.0	0.0	0.0	0.0	0.0	0.0

TABLE D-60

PDC CHILD RATING SCALE (ENGLISH ONLY RATED BY TEACHER)
RESPONSE DISTRIBUTION FOR: SECOND GRADE
TOTAL NUMBER OF RATINGS: 13

LENGTH OF TIME TEACHER HAD KNOWN CHILD:	PERCENT RESPONDING
LESS THAN MONTH	0.0
ONE-SIX MONTHS	0.0
SIX-TWELVE MONTHS	100.0
MORE THAN 12 MONTHS	0.0
HOW WELL TEACHER KNOWS CHILD	
NOT WELL	0.0
AVERAGE	92.3
BETTER THAN AVERAGE	7.7

I. SOCIAL COMPETENCE

	FREQUENTLY	MORE THAN AVERAGE	AVERAGE	LESS THAN AVERAGE	RARELY	NO RESPONSE
1.	0.0	30.8	53.8	7.7	7.7	0.0
2.	0.0	15.4	46.2	38.5	0.0	0.0
3.	0.0	15.4	61.5	15.4	7.7	0.0
4.	7.7	23.1	69.2	0.0	0.0	0.0
5.	7.7	7.7	38.5	46.2	0.0	0.0
6.	0.0	15.4	76.9	7.7	0.0	0.0
7.	0.0	15.4	69.2	15.4	0.0	0.0
8.	0.0	15.4	61.5	23.1	0.0	0.0
9.	0.0	30.8	61.5	7.7	0.0	0.0
10.	0.0	0.0	76.9	23.1	0.0	0.0
11.	0.0	7.7	53.8	38.5	0.0	0.0
12.	0.0	0.0	84.6	7.7	7.7	0.0
13.	0.0	7.7	46.2	38.5	7.7	0.0
14.	0.0	38.5	38.5	15.4	7.7	0.0
15.	0.0	7.7	69.2	23.1	0.0	0.0
16.	0.0	53.8	38.5	7.7	0.0	0.0
17.	0.0	7.7	76.9	15.4	0.0	0.0
18.	0.0	0.0	84.6	15.4	0.0	0.0
19.	0.0	15.4	61.5	23.1	0.0	0.0
20.	0.0	0.0	46.2	46.2	7.7	0.0
21.	0.0	0.0	84.6	15.4	0.0	0.0
22.	0.0	7.7	69.2	23.1	0.0	0.0
23.	0.0	53.8	46.2	0.0	0.0	0.0
24.	0.0	30.8	46.2	15.4	7.7	0.0

II. LEARNING HOW TO LEARN

	FREQUENTLY	MORE THAN AVERAGE	AVERAGE	LESS THAN AVERAGE	RARELY	NO RESPONSE
25.	7.7	7.7	84.6	0.0	0.0	0.0
26.	0.0	15.4	38.5	38.5	7.7	0.0
27.	7.7	38.5	30.8	23.1	0.0	0.0
28.	0.0	15.4	69.2	15.4	0.0	0.0
29.	0.0	0.0	30.8	53.8	15.4	0.0
30.	0.0	23.1	15.4	61.5	0.0	0.0
31.	0.0	23.1	30.8	38.5	7.7	0.0
32.	0.0	23.1	30.8	38.5	7.7	0.0
33.	0.0	30.8	69.2	0.0	0.0	0.0
34.	0.0	15.4	69.2	15.4	0.0	0.0
35.	0.0	23.1	46.2	30.8	0.0	0.0
36.	0.0	0.0	38.5	61.5	0.0	0.0
37.	0.0	23.1	61.5	15.4	0.0	0.0
38.	0.0	23.1	53.8	23.1	0.0	0.0
39.	100.0	0.0	0.0	0.0	0.0	0.0

TABLE D-61

PDC CHILD RATING SCALE (ENGLISH ONLY RATED BY TEACHER)
RESPONSE DISTRIBUTION FOR: THIRD GRADE, GEORGIA
TOTAL NUMBER OF RATINGS: 12

LENGTH OF TIME TEACHER HAD KNOWN CHILD:	PERCENT RESPONDING
LESS THAN MONTH	0.0
ONE-SIX MONTHS	16.7
SIX-TWELVE MONTHS	75.0
MORE THAN 12 MONTHS	8.3

HOW WELL TEACHER KNOWS CHILD	
NOT WELL	8.3
AVERAGE	83.3
BETTER THAN AVERAGE	0.0

I. SOCIAL COMPETENCE

	FREQUENTLY	MORE THAN AVERAGE	AVERAGE	LESS THAN AVERAGE	RARELY	NO RESPONSE
1.	8.3	33.3	33.3	8.3	16.7	0.0
2.	0.0	8.3	33.3	0.0	58.3	0.0
3.	16.7	8.3	58.3	16.7	0.0	0.0
4.	0.0	25.0	66.7	8.3	0.0	0.0
5.	0.0	16.7	50.0	16.7	16.7	0.0
6.	8.3	25.0	25.0	33.3	8.3	0.0
7.	8.3	25.0	41.7	25.0	0.0	0.0
8.	16.7	33.3	41.7	8.3	0.0	0.0
9.	33.3	16.7	33.3	16.7	0.0	0.0
10.	0.0	8.3	75.0	16.7	0.0	0.0
11.	0.0	8.3	41.7	33.3	16.7	0.0
12.	0.0	8.3	75.0	8.3	8.3	0.0
13.	0.0	8.3	41.7	16.7	33.3	0.0
14.	8.3	25.0	33.3	33.3	0.0	0.0
15.	8.3	33.3	58.3	0.0	0.0	0.0
16.	41.7	8.3	41.7	8.3	0.0	0.0
17.	8.3	8.3	66.7	16.7	0.0	0.0
18.	8.3	8.3	66.7	16.7	0.0	0.0
19.	16.7	16.7	58.3	8.3	0.0	0.0
20.	25.0	0.0	50.0	25.0	0.0	0.0
21.	8.3	8.3	66.7	16.7	0.0	0.0
22.	16.7	16.7	58.3	8.3	0.0	0.0
23.	25.0	0.0	58.3	16.7	0.0	0.0
24.	25.0	0.0	33.3	41.7	0.0	0.0

II. LEARNING HOW TO LEARN

	FREQUENTLY	MORE THAN AVERAGE	AVERAGE	LESS THAN AVERAGE	RARELY	NO RESPONSE
25.	8.3	16.7	66.7	8.3	0.0	0.0
26.	16.7	0.0	41.7	33.3	8.3	0.0
27.	8.3	8.3	50.0	16.7	16.7	0.0
28.	8.3	0.0	50.0	8.3	33.3	0.0
29.	0.0	0.0	16.7	16.7	66.7	0.0
30.	16.7	25.0	25.0	33.3	0.0	0.0
31.	0.0	8.3	50.0	16.7	25.0	0.0
32.	16.7	8.3	58.3	16.7	0.0	0.0
33.	16.7	16.7	66.7	0.0	0.0	0.0
34.	25.0	0.0	58.3	16.7	0.0	0.0
35.	0.0	16.7	0.0	25.0	58.3	0.0
36.	8.3	0.0	8.3	25.0	58.3	0.0
37.	8.3	25.0	41.7	25.0	0.0	0.0
38.	16.7	8.3	50.0	25.0	0.0	0.0
39.	100.0	0.0	0.0	0.0	0.0	0.0

TABLE D-62

FACES

RESPONSE DISTRIBUTION FOR: KINDERGARTEN

Variables are defined as follows:

801. Feelings about eating candy
802. Feelings about falling down
8801. Feelings about school this year
8802. Feelings about school next year
8803. Feelings about teacher
8804. Teachers feelings about child

VARIABLE	**N**	**MINIMUM**	**MAXIMUM**	**MEAN**	**STD DEV**
801.CANDY	24	2.0000	5.0000	3.6667	1.0072
802.FALL	24	1.0000	5.0000	2.0000	1.0632
8801.FSCHNOW	8	2.0000	5.0000	3.2500	1.0351
8802.FSCHNXY	8	2.0000	5.0000	3.7500	1.4880
8803.FCH.TO.T	8	2.5000	5.0000	4.0625	.82104
8804.FT.TO.CH	8	1.5000	5.0000	4.0000	1.1952

TABLE D-63

FACES

RESPONSE DISTRIBUTION FOR: FIRST GRADE

Variables are defined as follows:

801. Feelings about eating candy
802. Feelings about falling down
8801. Feelings about school this year
8802. Feelings about school next year
8803. Feelings about teacher
8804. Teachers feelings about child

VARIABLE	N	MINIMUM	MAXIMUM	MEAN	STD DEV
801.CANDY	27	2.0000	5.0000	4.4815	.75296
802.FALL	27	1.0000	3.0000	1.4074	.63605
8801.FSCHNOW	23	1.0000	5.0000	4.1304	1.0576
8802.FSCHNXY	23	1.0000	5.0000	3.7391	1.4528
8803.FCH.TO.T	23	2.0000	5.0000	4.5072	.90938
8804.FT.TO.CH	23	1.0000	5.0000	4.2754	1.0379

TABLE D-64

FACES

RESPONSE DISTRIBUTION FOR: SECOND GRADE

Variables are defined as follows:

801. Feelings about eating candy
802. Feelings about falling down
8801. Feelings about school this year
8802. Feelings about school next year
8803. Feelings about teacher
8804. Teachers feelings about child

VARIABLE	N	MINIMUM	MAXIMUM	MEAN	STD DEV
801.CANDY	27	3.0000	5.0000	4.4815	.75296
802.FALL	27	1.0000	3.0000	1.5556	.64051
8801.FSCHNOW	21	1.0000	5.0000	4.0000	1.0954
8802.FSCHNXY	21	1.0000	5.0000	3.4286	1.5353
8803.FCH.TO.T	21	2.0000	5.0000	4.2937	.93676
8804.FT.TO.CH	21	1.0000	5.0000	4.0714	1.1650

TABLE D-65

FACES

RESPONSE DISTRIBUTION FOR: THIRD GRADE, GEORGIA & MARYLAND

Variables are defined as follows:

801. Feelings about eating candy
802. Feelings about falling down
8801. Feelings about school this year
8802. Feelings about school next year
8803. Feelings about teacher
8804. Teachers feelings about child

VARIABLE	N	MINIMUM	MAXIMUM	MEAN	STD DEV
801.CANDY	54	3.0000	5.0000	4.7222	.52903
802.FALL	54	1.0000	3.0000	1.4630	.66483
8801.FSCHNOW	47	1.0000	5.0000	3.7660	1.2017
8802.FSCHNXY	47	1.0000	5.0000	3.8085	1.2272
8803.FCH.TO.T	47	1.0000	5.0000	4.4220	.91079
8804.FT.TO.CH	47	2.0000	5.0000	4.2766	.75604

TABLE D-66

PRESCHOOL INTERPERSONAL PROBLEM SOLVING TEST

RESPONSE DISTRIBUTION FOR: KINDERGARTEN

VARIABLE	N	MINIMUM	MAXIMUM	MEAN	STD DEV
SOLUTNS	24	0.	6.0000	3.7500	1.6746
REL.GL.%	24	0.	1.0000	.20833	.41485
REL.GOAL	5	1.0000	7.0000	2.2000	2.6833
SUB.GL.%	24	0.	1.0000	.20833	.41485
SUB.GOAL	5	1.0000	5.0000	2.2000	1.7889
IRREL%	24	0.	1.0000	.25000	.44233
IRREL	6	1.0000	7.0000	2.1667	2.4014
DKRNR%	24	0.	1.0000	.70833	.46431
DKRNR	17	1.0000	11.000	4.2941	2.7332
NPHYSOL	24	0.	6.0000	2.8750	1.6235
ASK%	24	0.	1.0000	.66667	.48154
PLEASE%	24	0.	1.0000	.37500	.49454
LOAN%	24	0.	1.0000	.41667 -1	.20412
FAIR%	24	0.	1.0000	.54167	.50898
TRADE%	24	0.	1.0000	.16667	.38069
AUTHINT%	24	0.	1.0000	.50000	.51075
TRICK%	24	0.	1.0000	.83333 -1	.28233
FINAGLE%	24	0.	0.	0.	
MANAFF%	24	0.	1.0000	.16667	.38069
MAD %	24	0.	0.	0.	
WAIT%	24	0.	1.0000	.20833	.41485
FUTURE%	24	0.	1.0000	.12500	.33783
PHYSOL	24	0.	2.0000	.87500	.79741
FORCE%	24	0.	1.0000	.37500	.49454
PHYS%	24	0.	1.0000	.33333	.48154
DAMAGE %	24	0.	0.	0.	
COMMAND%	24	0.	1.0000	.16667	.38069

TABLE D-67

PRESCHOOL INTERPERSONAL PROBLEM SOLVING TEST

RESPONSE DISTRIBUTION FOR: FIRST GRADE

VARIABLE	N	MINIMUM	MAXIMUM	MEAN	STD DEV
SOLUTNS	27	0.	6.0000	3.3333	1.5933
REL.GL.%	27	0.	1.0000	.11111	.32026
REL.GOAL	3	1.0000	2.0000	1.6667	.57735
SUB.GL.%	27	0.	1.0000	.74074 -1	.26688
SUB.GOAL	2	1.0000	3.0000	2.0000	1.4142
IRREL%	27	0.	1.0000	.29630	.46532
IRREL	8	1.0000	3.0000	1.6250	.74402
DKRNR%	27	0.	1.0000	.92593	.26688
DKRNR	25	1.0000	8.0000	3.3600	1.5513
NPHYSOL	27	0.	6.0000	2.6667	1.5191
ASK%	27	0.	1.0000	.70370	.46532
PLEASE%	27	0.	1.0000	.37037	.49210
LOAN%	27	0.	1.0000	.11111	.32026
FAIR%	27	0.	1.0000	.48148	.50918
TRADE%	27	0.	1.0000	.18519	.39585
AUTHINT%	27	0.	1.0000	.25926	.44658
TRICK%	27	0.	0.	0.	
FINAGLE%	27	0.	1.0000	.37037 -1	.19245
MANAFF%	27	0.	1.0000	.37037 -1	.19245
MAD %	27	0.	0.	0.	
WAIT%	27	0.	1.0000	.29630	.46532
FUTURE%	27	0.	1.0000	.18519	.39585
PHYSOL	27	0.	2.0000	.66667	.83205
FORCE%	27	0.	1.0000	.37037	.49210
PHYS%	27	0.	1.0000	.18519	.39585
DAMAGE %	27	0.	0.	0.	
COMMAND%	27	0.	1.0000	.11111	.32026

TABLE D-68

PRESCHOOL INTERPERSONAL PROBLEM SOLVING TEST

RESPONSE DISTRIBUTION FOR: SECOND GRADE

VARIABLE	N	MINIMUM	MAXIMUM	MEAN	STD DEV
SOLUTNS	27	1.0000	8.0000	4.8889	1.8257
REL.GL.%	27	0.	1.0000	.22222	.42366
REL.GOAL	6	1.0000	3.0000	1.5000	.83666
SUB.GL.%	27	0.	1.0000	.11111	.32026
SUB.GOAL	3	1.0000	4.0000	2.0000	1.7321
IRREL%	27	0.	1.0000	.18519	.39585
IRREL	5	1.0000	4.0000	2.0000	1.2247
DKBNR%	27	0.	1.0000	.88889	.32026
DKBNR	24	1.0000	9.0000	4.0000	2.3406
NPHYSOL	27	1.0000	6.0000	4.0741	1.3847
ASK%	27	0.	1.0000	.92593	.26688
PLEASE%	27	0.	1.0000	.37037	.49210
LOAN%	27	0.	1.0000	.14815	.36201
FAIR%	27	0.	1.0000	.70370	.46532
TRADE%	27	0.	1.0000	.29630	.46532
AUTHINT%	27	0.	1.0000	.48148	.50918
TRICK%	27	0.	1.0000	.22222	.42366
FINAGLE%	27	0.	1.0000	.37037 -1	.19245
MANAFF%	27	0.	1.0000	.18519	.39585
MAD %	27	0.	0.	0.	
WAIT%	27	0.	1.0000	.44444	.50637
FUTURE%	27	0.	1.0000	.25926	.44658
PHYSOL	27	0.	3.0000	.81481	.96225
FORCE%	27	0.	1.0000	.33333	.48038
PHYS%	27	0.	1.0000	.33333	.48038
DAMAGE %	27	0.	0.	0.	
COMMAND%	27	0.	1.0000	.14815	.36201

TABLE D-69

PRESCHOOL INTERPERSONAL PROBLEM SOLVING TEST

RESPONSE DISTRIBUTION FOR: THIRD GRADE, GEORGIA & MARYLAND

VARIABLE	N	MINIMUM	MAXIMUM	MEAN	STD DEV
SOLUTNS	54	1.0000	8.0000	5.0185	1.6193
REL.GL.%	54	0.	1.0000	.37037	.48744
REL.GOAL	20	1.0000	3.0000	1.4500	.75915
SUB.GL.%	54	0.	1.0000	.12963	.33905
SUB.GOAL	7	1.0000	2.0000	1.1429	.37796
IRREL%	54	0.	1.0000	.29630	.46091
IRREL	16	1.0000	2.0000	1.0625	.25000
DKRNR%	54	0.	1.0000	.70370	.46091
DKRNR	38	1.0000	9.0000	3.1842	2.0906
NPHYSOL	54	1.0000	7.0000	4.1296	1.3465
ASK%	54	0.	1.0000	.90741	.29258
PLEASE%	54	0.	1.0000	.40741	.49597
LOAN%	54	0.	1.0000	.11111	.31722
FAIR%	54	0.	1.0000	.81481	.39210
TRADE%	54	0.	1.0000	.44444	.50157
AUTHINT%	54	0.	1.0000	.38889	.49208
TRICK%	54	0.	1.0000	.20370	.40653
FINAGLE%	54	0.	1.0000	.11111	.31722
MANAFF%	54	0.	1.0000	.92593 -1	.29258
MAD %	54	0.	1.0000	.18519 -1	.13608
WAIT%	54	0.	1.0000	.46296	.50331
FUTURE%	54	0.	1.0000	.16667	.37618
PHYSOL	54	0.	3.0000	.88889	1.0581
FORCE%	54	0.	1.0000	.35185	.48203
PHYS%	54	0.	1.0000	.33333	.47583
DAMAGE %	54	0.	0.	0.	
COMMAND%	54	0.	1.0000	.20370	.40653

APPENDIX E

PDC Observation Behavior Categories

and

Relative Frequencies of Behaviors Observed Using PDC Observation System

PDC Observation System: Definition and Examples of Behavior Categories

Involvement (Categories 1 and 2)

Category 1. Noninvolved (NONINV)

Category 1 is coded when the child is not interacting with a peer, adult, or object and is not doing anything else that seems to have a purpose. This category includes instances where the child is looking at a person, but the person is not looking at or talking directly to the child. In other words, it is not a reciprocal interaction; the child is merely watching someone who is not paying attention to him. When this category is coded, you do not code categories 2 through 4. You do code category 5.

Examples of this behavior category:

- The child is staring into space, not paying attention to the things going on around him.
- The child is aimlessly wandering around the classroom and doesn't seem to have any purpose or intention in mind.
- The child is holding an object, such as a block, but doesn't do anything with it. And he doesn't seem to be planning to do anything with it.
- The child is watching two boys build a block tower. The boys are not looking at the target child.
- The child is watching the aide who is helping another child make a kite. The teacher is directing her attention only on the child making the kite.

Category 2. Involved

Category 2 is coded when the target child is interacting with a peer(s), adult(s), or object(s) or is involved in an observable directed behavior, such as singing to himself. If the child is involved in any of these ways during the 5-second observation interval, his behavior is coded in two subcategories to indicate: (a) the context of the activity (for example, interacting with people and/or materials); and (b) the language spoken during the activity, if any. In order to code category 2, you place a slash mark in each of these subcategories (2a and 2b) to describe the child's involvement. After coding the subcategories of category 2, you will look at categories 3 and 4 and mark any that further describe the child's behavior.

Subcategories. When you have decided that category 2 applies to the child's behavior, you must place a slash mark on one item in each of two subcategories: 2a and 2b.

Subcategory 2a. Focus of child's attention. If the child is "involved" in an activity, then you must decide which of the following two items describe the focus of the child's attention: social or non-social.

Social (SOC): The child is paying attention to another person (peer or adult) by looking at or listening to this person. The person the child is looking at or listening to must also be looking at the child or speaking directly to the child, either as an individual or as a member of a group. A child may also (in addition to looking at or listening to) be paying attention to another person by sharing materials and working on a common project, talking to and/or touching the person. This item is coded when the child interacts with both persons and objects (either at the same time or one following the other) during the 5 second interval.

Examples:

- The child is sitting on the teacher's lap listening to a story and is helping tell parts of the story.
- The child is calling a peer a name.
- The child is listening to a peer who is telling him how to paint his picture.
- The child is playing a lotto game with the teacher.
- The child and a peer are looking at each other as they eat their snack.

Note: If the child is social with a peer, adult, or both peer and adult, you must code the appropriate items in categories 3 and/or 4 that best describe the child's social interactions.

Nonsocial (NSOC): The child is paying attention to objects by looking at and/or touching those objects or the child is engaged in some other observable, directed behavior which does not involve other persons (such as singing to himself).

Examples:

- The child is quietly putting a puzzle together at the toy table without talking to anyone else.
- The child is skipping alone around the room.
- The child is sitting on the rug singing to himself.

Note: When this item is marked you skip categories 3 and 4 and code category 5.

Subcategory 2b. Language spoken during the activity. If the child is "involved" in an activity then you must decide which one of the following four items describes the behavior: verbal in English, verbal in Spanish, verbal in combined English and Spanish, or nonverbal.

Verbal in English (VENG): While engaging in activities with people and/or objects, the child speaks only in English.

Verbal in Spanish (VSP): While engaging in activities with people and/or objects, the child speaks only in Spanish.

Verbal in Combined English and Spanish (VCOMB): While engaging in activities with people and/or objects, the child speaks a combination of English and Spanish or uses Spanglish.

Nonverbal (NONV): While engaging in activities with people and/or objects, the child does not speak.

Category 3. Interactions with Peer

Category 3 is coded when the target child interacts with a peer(s) by looking at, listening to, talking with, or sharing materials and working on a common project. The peer the child is looking at or listening to must be looking at the child or speaking directly to the target child. If the child is interacting with a peer in any of these ways during the 5-second interval, his behavior is coded in four subcategories to indicate: (a) the type of peer interaction; (b) the nature of the peer interaction; (c) the purpose of the peer interaction; and (d) the role played during the peer interaction. In order to code category 3, you place a slash mark in each of these subcategories (3a, 3b, 3c, and 3d) to describe the child's interaction with a peer(s). If more than one of the items in a given subcategory occurs during the 5-second interval, put a slash mark beside the item that occurred last.

Subcategories. When you have decided that category 3 applies to the child's behavior, you must place a slash mark on one item in each of the four subcategories: 3a, 3b, 3c, and 3d.

Subcategory 3a. Type of peer interaction. If the child is interacting with a peer, then you must decide which one of the following three items best describes the behavior: controlling, assertion, or other.

Controlling (CNT): The child attempts verbally and/or nonverbally (i.e., physical gestures) to influence or get the attention of a peer.

Examples:

- The child approaches a peer showing him a game and says, "Let's play this game."
- The child goes over to a peer and grasps his hand, then leads him to aquarium to view the fish.
- The child yells "Hey John!" across the room. John turns and looks over to the child.

- The child pushes a peer away from the drinking fountain, knocking her down onto the floor.

Assertion (AST): The child resists verbal or nonverbal peer attempts to influence his behavior or direct him to do something.

Examples:

- When asked by a peer to join in an art activity, the child says amiably, "I want to go play with the blocks."
- When the child attempts to join a group of peers building a block structure, one of the peers says, "You can't play here." The child responds by knocking down the block structure and building another one in its place.
- When a peer comes over and asks the child to come into the house, the child ignores the peer's question and does not respond.
- When asked by a peer to play a game, the child responds, "It's my turn to paint now" and pushes the peer away.

Other (OTH): This item is marked for each interaction with a peer that is clearly not an attempt to control or assertion.

- The child argues with another child about whose turn it is to ride the tricycle and calls the child several names. The peer starts to cry.
- For no apparent reason, the child hits another child in the stomach.
- Two children are talking.
- While playing with playdough at the art table, the child divides and shares his playdough with a peer who does not have any.

Subcategory 3b. Nature of peer interactions. If the child is interacting with a peer, then you must decide which of the following two items best describe the behavior: negative or positive/neutral.

Negative (NEG): The child expresses verbal and/or nonverbal agression or hostility toward the person he is interacting with.

Examples:

- When the child attempts to join a group of peers building a block structure, one of the peers says, "You can't play here." The child responds by knocking down the block structure and building another in its place.
- The peer pushes a peer away from the drinking fountain knocking her down onto the floor.
- When asked by a peer to play a game, the child responds, "It's my turn to paint now" and pushes the peer away.
- For no apparent reason, the child hits another child in the stomach.

Positive/Neutral (POS/NT): This item is marked for each behavior that is clearly not negative.

- The child is playing a lotto game with a peer.
- When asked by a peer to join in an art activity, the child says amiably, "I want to go play with the blocks."
- While playing with playdough at the art table, the child divides and shares his playdough with a peer who does not have any.
- The child approaches a peer showing him a game and says, "Let's play this game."

Subcategory 3c. Purpose of peer interation. If the child is interacting with a peer, then you must decide which one of the following four items best describe the purpose of the child's interaction: information, assistance/materials, support, or nonapplicable.

Information (INFO): The child requests or provides factual statements or explanations concerning a task, a problem, a causal relationship or other events and situations in his environment.

Examples:

- While playing with a peer at the workbench, the child asks, "Where's the hammer?"
- The child asks a peer, "Why are you doing that?"
- The child turns to a nearby peer and says, "Look its snowing outside."
- When asked a question by a peer, the child simply replies, "No."

Assistance/Materials (AS/MAT): The child requests or provides physical assistance or materials.

Examples:

- The child requests a peer to come over and help obtain a toy from another peer.
- At snack time the child helps a peer pour her juice.
- The child hands a peer a block.
- The child tells a peer, "Make a "J" for me."

Support (SUP): The child is used by or uses a peer for obtaining comfort, protection, and/or reassurance after a hurt, disappointment, or other problem situations. The child does not seek or provide assistance or information for solving the problem.

Examples:

- The child puts her arms around a crying peer.
- The child takes a hold of a peer's hand after he has been slugged by another child.

- The child says to another child, "Bob took my truck." The child does not ask the peer to help get it back.

- After falling off a tricycle, the child cries to a peer, "My knee hurts."

Nonapplicable (NA): This item is marked whenever the purpose of the child's interaction with a peer is clearly not one of requesting or providing information, assistance, materials, or emotional support.

Examples:

- The child and a peer are painting a monster picture together at the art easel.

- A peer tells the child not to do something.

- The child laughs at a peer's antics. The peer looks over and giggles.

- A peer tells the child, "I ate lunch at a restaurant."

Subcategory 3d. Role played during peer interactions. If the child is interacting with a peer, then you must decide which one of the following three items best describe the child's role: requester, giver, or nonapplicable.

Requester (REQ): The child requests information, assistance, materials, or emotional support by posing a question, making a demand, or in some manner indicating (i.e., physical gestures) a need of help, materials, information, or emotional support from a peer.

- The child asks a peer, "Why are you doing that?"

- The child asks a peer, "How do you make a "J"?"

- The child holds hands with a friend immediately after the teacher has scolded the child.

Giver (GIVE): The child gives information, assistance, materials, or emotional support to a peer in the form of factual statements, explanations, or physical gestures. The child may provide this spontaneously or on the peer's request.

- The child turns to a nearby peer and says, "Look its snowing outside!"
- The child hands a peer a block.
- The child puts her arms around a crying peer.
- The child shows a peer how to mix blue and green paint together to make a new color.

Nonapplicable (NA): The child is not requesting or providing information, assistance, materials, or emotional support. This item is marked whenever non-applicable is coded in subcategory 3c.

Examples:

- The child listens to a peer tell a story. The child did not ask the peer to tell the story.
- The child and a peer are painting a monster picture together at the art easel.
- A peer comes over to the child and says, "Its clean-up time."
- The child pushes a peer away from the drinking fountain, knocking her down onto the floor.

Category 4. Interactions with Adult

Category 4 is coded when the target child interacts with an adult(s) by looking at, listening to, talking with, or sharing materials and working on a common project. The adult the child is looking at or listening to must be looking at the child or speaking directly to the target child. If the child is interacting with an adult in any of these ways during the 5-second interval, his behavior is coded in four subcategories to indicate: (a) the type of adult interaction; (b) the nature of the adult interaction; (c) the purpose of the adult interaction; and (d) the role played during the adult interaction. In order to code category 4, you place a slash mark in each of these subcategories (4a, 4b, 4c, and 4d) to describe the child's interaction with an adult(s). If more than one of the items in a given subcategory occurs during the 5-second interval, put a slash mark beside the item that occurred last.

Subcategories. When you have decided that category 4 applies to the child's behavior, you must place a slash mark on one item in each of the four subcategories: 4a, 4b, 4c, and 4d.

Subcategory 4a. Type of adult interaction. If the child is interacting with an adult, then you must decide which one of the following three items best describes the type of adult interaction: controlling, assertion, or other.

Controlling (CNT): The child attempts verbally and/or nonverbally (i.e., physical gestures) to influence or get the attention of an adult.

Examples:

- The child tells the teacher, "Shut-up."
- The child tells an aide, "Leave me alone," and pushes her away.
- The child approaches an aide and says, "I want you to build this block tower with me."
- The child goes over to a teacher, grasps her hand then leads her to the art table and asks, "Do you know what I made?"

Assertion (AST): The child resists verbal and nonverbal adult attempts to influence his behavior or direct him to do something.

Examples:

- When asked to clean-up his desk by a volunteer aide the child replys, "I don't have to," and kicks the aide in the leg.
- When the teacher asks the child to finish his snack, the child throws the cookies on the floor, dumping over the milk.

- When the teacher says "clean-up time," the child tells the teacher, "I've cleaned up my share; I can go to snack now."

- When told by the teacher to go to the art area, the child asserts, "I'd rather go play at the toy table."

- The teacher addresses a question to the child and the child ignores her.

Other (OTH): This item is marked for each interaction with an adult that is clearly not an attempt to control or assertion.

Examples:

- The child hits a parent aide.

- The child listens to the teacher explain a game.

- When the teacher asks the child to share the paste jar with another peer, the child moves the paste to the middle of the table for sharing.

- The child and the teacher are talking as they both carry the snacks to the table.

Subcategory 4b. Nature of adult interactions. If the child is interacting with an adult, then you must decide which of the following two items best describe the nature of the behavior: negative or positive/neutral.

Negative (NEG): The child expresses verbal and/or nonverbal agression or hostility toward the adult he is interacting with.

Examples:

- The child screams to the adult, "Shut-up!"

- The child hits the adult.

- The child tells an aide, "Leave me alone," and pushes her away.

- When the teacher asks the child to finish his snack, the child throws the cookies on the floor, dumping over the milk.

Positive/Neutral (POS/NT): This item is marked for each behavior that is clearly not negative.

Examples:

- The child approaches an aide and says, "I want you to build this block tower with me."
- The teacher addresses a question to the child. The child does not respond.
- The child approaches the teacher and offers to help put away the dress-up clothes.
- The child listens to the teacher explain a game.

Subcategory 4c. Purpose of adult interaction. If the child is interacting with an adult, then you must decide which one of the following four items best describe the purpose of the child's interaction: information, assistance/materials, support, or nonapplicable.

Information (INFO): The child requests or provides factual statements or explanations concerning a task, a problem, a causal relationship or other events and situations in his environment.

Examples:

- When asked by the teacher to tell what shape she was holding in her hand, the child responds, "That's a circle."
- The child tells the teacher that today is his birthday.
- The child asks a parent aide where the paint brushes are.
- The child asks the teacher, "When is it going to be lunch time?"

Assistance/Materials (AS/MAT): The child requests or provides physical assistance or materials.

Examples:

- After an unsuccessful attempt to play with playdough, the child goes over to the teacher and demands, "Make her share with me."
- After using the toilet, the child goes over to the teacher and waits for him to zip up her pants.
- The child helps the aide put the blocks on the shelf.
- The child hands the cookie tray to the teacher.

Support (SUP): The child is used by or uses an adult for obtaining comfort, protection, and/or reassurance after a hurt, disappointment, or other problem situations. The child does not seek or provide assistance or information for solving the problem.

Examples:

- The child holds hands with an adult immediately after a peer has hit her.
- During a loud rainstorm, the child goes over to the teacher and sits in her lap.
- The child rubs the teacher's head after she bumped it on a cupboard door.
- The child says to a parent aide, "I'm sorry you got a bad cold."

Nonapplicable (NA): This item is marked whenever the purpose of the child's interaction with an adult is clearly not one of requesting or providing information, assistance, materials, or emotional support.

Examples:

- When the teacher asks the child to finish his snack, the child throws the cookies on the floor, dumping over the milk.
- The child listens to a volunteer parent aide tell a story.
- The child and the teacher smile at each other as they eat their snack.
- The child hits an aide.

Subcategory 4d. Role played during adult interactions. If the child is interacting with an adult, then you must decide which one of the following three items best describe the child's role: requester, giver, or nonapplicable.

Requester (REQ): The child requests information, assistance, materials, or emotional support by posing a question, making a demand, or in some manner indicating (i.e., physical gestures) a need of help, materials, information, or emotional support from an adult.

Examples:

- After an unsuccessful attempt to play with the playdough, the child goes over to the teacher and demands, "Make her share with me."
- The child holds hands with an adult immediately after a peer has hit her.
- The child asks a parent aide where the paint brushes are.
- The child asks the teacher, "When is it going to be lunch time?"

Giver (GIVE): The child gives information, assistance, materials, or emotional support to a peer in the form of factual statements, explanations, or physical gestures. The child may provide this spontaneously or on the adult's request.

Examples:

- The child tells the teacher that today is his birthday.
- The child answers the aide's question.
- The child rubs the teacher's head after she bumped it on a cupboard door.
- The child hands an aide her coat.

Nonapplicable (NA): The child is not requesting or providing information, assistance, materials, or emotional support. This item is marked whenever nonapplicable is coded in subcategory 4c.

- The teacher tells the child to clean up his toys.
- The child listens to the teacher explain a game.
- The child listens to the teacher explain how to do a problem.
- The volunteer aide gives the child a hug.
- The child bites the teacher.

Category 5. Classroom Interaction Capacity

The kinds of child-child and child-adult interactions that are likely to happen vary according to the time of day or activity schedule. So it is necessary for the observer to describe the classroom's interaction capacity during each observation unit. To do this, the observer surveys the classroom in order to determine the degree of interactions occurring among children and adults after each 5-second observation. The observer's attention is no longer directed toward the target child but on the classroom as a whole. Thus, even though the target child's behavior is not congruent with the behavior of the other children (the focal child is being restricted by an adult), the observer still indicates to what extent interactions may occur during that observational unit.

For category 5, the observer looks at the whole classroom and indicates the interaction capacity of the classroom during each observation unit by placing a slash mark next to the appropriate item. The items are: maximal, moderate and minimal.

Maximal: (MAX) This item refers to those parts of the day in which children and adults are free to initiate or maintain spontaneous interactions (verbally or physically) among themselves. The children are generally able to choose their own behavior, with minimal structuring/direction by an adult. These periods may be labelled by the teacher as "free play" or "free choice" periods.

Examples:

- For a thirty-minute period, children are free to choose and move at their own discretion among any activity available in the room.
- For a fifteen-minute period, children are free to choose and move at their own discretion among activities prepared and set up by an adult. Activities are carried out individually by the children with minimal direction by the adult.

Moderate: (MOD) This item refers to those parts of the day in which the opportunity for spontaneous interaction among adults and children is substantially reduced. During this period, classroom behavior is typically less decided by children and more directed by an adult. There is still some opportunity for spontaneous interactions to occur within this given structure.

Examples:

- For a ten-minute period, the children are assigned to a small group (3-8 children) where the teacher is reading a story. The children are expected to remain with the group, but may talk to the teacher and other children at points during the story or after the story is completed.
- After a thirty-minute period of "free play", the children are engaged in "clean-up" activities. The children may interact with one another, but they are all expected to help the teacher or sit in a designated area of the room.

- During snack and meal times, the children are assigned to a given table and adult. The children are expected to remain at the table, but may interact among themselves or with their assigned adult.

Minimal: This item refers to those parts of the day in which children are not free to initiate/maintain spontaneous interactions (verbally or physically) among themselves. Classroom behavior of the children is primarily controlled and directed by an adult.

Examples:

- During a 15-minute period, all the children are sitting in a large group listening to the teacher's story. The children are expected to pay attention and not interact (physically or verbally) during this period.

- During snack and meal times, the children are assigned to tables and an adult. The children are expected to use this time for eating, not interacting.

- During a 20-minute period, all the children are sitting in a large group singing with the teacher. All children are expected to participate and interactions among children are discouraged and restricted.

Relative Frequencies of Behaviors Observed Using PDC Observation System

Table E-1

Means and Standard Deviations of Child-Peer Interactions[1]

	ACROSS ALL activity levels *(N=441)*		MAXIMUM activity levels *(N=270)*		MODERATE activity levels *(N=300)*		MINIMUM activity levels *(N=86)*	
Observation Variable	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Description of Child-Peer Interaction								
Negative Control	( )[2]	--	.01	.03	.01	.06	0	--
Positive Control	.28	.27	.33	.31	.22	.29	.25	.33
Negative Assert	( )	--	.01	.03	( )	--	0	--
Positive Assert	.06	.13	.07	.17	.10	.15	.07	.21
Negative Other	( )	--	( )	--	( )	--	.01	.11
Positive Other	.50	.55	.59	.34	.71	.32	.67	.38
Purpose of Child-Peer Interaction								
Request Information	.08	.14	.08	.16	.08	.16	.09	.20
Give Information	.45	.28	.44	.30	.47	.34	.45	.38
Request Assistance/Materials	.15	.21	.19	.26	.13	.25	.13	.25
Give Assistance/Materials	.05	.12	.07	.16	.04	.13	.04	.13
Request Support	( )	--	( )	--	( )	--	( )	--
Give Support	( )	--	( )	--	( )	--	.01	.11

---

[1] Estimates represent relative frequencies.

[2] Relative frequency of this category fell between .00 and .01.

Table E-2

Means and Standard Deviations of Child-Adult Interaction Variables[1]

Spring 1976

	ACROSS ALL activity levels *(N=377)*		MAXIMUM activity levels *(N=169)*		MODERATE activity levels *(N=97)*		MINIMUM activity levels *(N=98)*	
Observation Variable	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Description of Child-Peer Interaction								
Negative Control	( )[2]	--	( )	--	0	--	0	--
Positive Control	.30	.32	.40	.36	.26	.33	.21	.34
Negative Assert	0	0	0	--	0	--	0	0
Positive Assert	.03	.11	.03	.10	.03	.12	.03	.03
Negative Other	( )	--	( )	.02	( )	--	0	0
Positive Other	.67	.32	.57	.37	.71	.35	.77	.37
Purpose of Child-Peer Interaction								
Request Information	.01	.11	.14	.26	.01	.01	.11	.27
Give Information	.20	.35	.47	.39	.28	.41	.53	.43
Request Assistance/Materials	.02	.10	.11	.23	.01	.10	.05	.20
Give Assistance/Materials	( )	--	.06	.18	0	--	.05	.18
Request Support	0	--	.02	.11	0	--	.01	.10
Give Support	0	--	0	0	0	--	0	0

---

[1] Estimates represent relative frequencies.

[2] Relative frequency of this category fell between .00 and .01.

Table E-3

Means and Standard Deviations of Children's Involvement and Verbal Behavior[1]

Spring 1976

	ACROSS ALL activity levels *(N=469)*		MAXIMUM activity levels *(N=318)*		MODERATE activity levels *(N=367)*		MINIMUM activity levels *(N=185)*	
Observation Variable	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Children's Involvement								
Noninvolved	.00	.13	.08	.15	.08	.13	.08	.13
Social	.52	.20	.47	.25	.52	.24	.61	.28
Nonsocial	.40	.19	.45	.25	.40	.24	.31	.28
Children's Verbal Behavior								
Verbal English	.31	.81	.32	.22	.30	.21	.29	.24
Verbal Spanish	.03	.09	.04	.13	.03	.09	( )	--
Verbal Combination	( )[2]	--	( )	--	( )	--	.00	.00
Nonverbal	.66	.16	.64	.21	.66	.20	.72	.24

---

[1] Figures represent relative frequencies.

[2] Relative frequency of this category fell between .00 and .01.

APPENDIX F

Center and School Characteristics

These data, collected in fall 1975, are presented here for the record. They may prove useful in the future for the purpose of explaining observed differences in performance, or simply for the purpose of describing centers and schools (in terms of size, adult-child ratio, etc.). The arrows at the top of each form indicate Head Start-to-elementary school feeding patterns.

Table F-1

Visalia, California

PDC		Comparison	
Head Start Centers	Elementary Schools	Head Start Centers	Elementary Schools
Visalia →	Crowley	Goshen →	Goshen
		Linnell Camp →	Union

All Centers and Schools Combined, by Treatment

	Total classes		Total children		Total teachers				Total paid aides				Child:tchr ratio				Child:aide ratio			
					full-time		part-time		full-time		part-time		full-time		part-time		full-time		part-time	
	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com
HS	4	3	45	45	5	4	0	0	10	6	1	0	9	11			5	8	45	
K	2	3	65	99	2	3	0	0	3	*	1	0	33	33			22	*	65	
1	3	3	80	85	3	3	0	0	3	*	0	0	27	28			27	*		
2	2	3	48	88	2	3	0	0	3	*	0	0	24	29			16	*		
3	2	3	61	90	2	3	0	0	3	*	0	0	31	30			20	*		
Multi-level	0	2	0	60	0	2	0	0	0	0	0	0		30				*		
K-3 sum	9	14	254	422	9	14	0	0	12	32	1	0								
K-3 class average			28	30	1	1	0	0	1.3	2.3	0.1	0								
K-3 combined child/adult average													24	23.5			21	13	254	

*Aides serve in floating "teams" from K to 3.

Table F-2

Pueblo, Colorado

PDC			Comparison		
Head Start Centers		Elementary Schools	Head Start Centers		Elementary Schools
Fountain	→	Fountain	Bambi	→	Beulah Heights
Irving	→	Irving	West 11th	→	Thatcher

All Centers and Schools Combined, by Treatment

	Total classes		Total children		Total teachers full-time		Total teachers part-time		Total paid aides full-time		Total paid aides part-time		Child:tchr ratio full-time		Child:tchr ratio part-time		Child:aide ratio full-time		Child:aide ratio part-time	
	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com
HS	4	3	62	61	2	2		2			4	4	31	30.5		30.5			15.5	15.2
K	5	7	117	164	2	3	7	1	2		1		58.5	54.6	16.7	164	58.5		1	
1	6	6	104	157	6	6	6		1		1		17.3	26.1	17.3		104		1	
2	4	6	74	135	4	6			1		0		18.5	22.5			74		0	
3	5	5.5*	77	118	5	6			1		0		15.4	0.9			77		0	
Multi-level																				
K-3 sum	24	27.5	434	635	17	21	13	1	5		6	4								
K-3 class average			18	23	0.7	0.8	0.5	0.04	0.2		0.3	0.2								
K-3 combined child/adult average													25.5	30.2	33.3	635	86.8		105.8	

238 *.5 represents a third and fourth grade combination class.

Table F-3

Norwalk, Connecticut

PDC Head Start Centers	PDC Elementary Schools	Comparison Head Start Centers	Comparison Elementary Schools
Broad River Nathaniel Ely	→ Jefferson Magrath Columbus	Broad River Nathaniel Ely	→ Brookside Cranbury Naramake Silvermine Wolfpit

All Centers and Schools Combined, by Treatment

	Total classes		Total children		Total teachers				Total paid aides				Child:tchr ratio				Child:aide ratio			
					full-time		part-time		full-time		part-time		full-time		part-time		full-time		part-time	
	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com
HS	8	8	159	159	16	16			1	1			9.9	9.9			159	159		
K	8.5	15	237	382	5	8.5		1	2		1		47.4	44.9		382				
1	10	14.5	216	359	10	13.5			3			1	21.6	26.5						
2	9	14.5	219	365	9	14.5			1	1	1	1	24.3	25.1						
3	9	15	214	382	9	15					1	1	23.7	25.4						
Multi-level		2		42		2								21						
K-3 sum	36.5	61	886	1530	33	53.5		1	6	1	3	3								
K-3 class average			24.2	25	0.9	0.8		0.01	0.2	0.01	0.08	0.04								
K-3 combined child/adult average													26.8	28.5		1530				

Table F-4

Pahokee, Florida

PDC		Comparison	
Head Start Centers	Elementary Schools	Head Start Centers	Elementary Schools
Fremd Village →	Pahokee	Washington Park →	Rosenwald
Okeechobee →	Gove	Lake Harbor →	Moorehaven

All Centers and Schools Combined, by Treatment

	Total classes		Total children		Total teachers				Total paid aides				Child:tchr ratio				Child:aide ratio			
					full-time		part-time		full-time		part-time		full-time		part-time		full-time		part-time	
	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com
HS	4	6	77	75	2	3	0	0	11	14	0	0	39	25			7	*		*
K	8	*	207	*	9	*	0	*	10	*	0	*	26	*		*	21	*		*
1	10	*	243	*	11	*	0	*	3	*	0	*	24	*		*	81	*		*
2	10	*	246	*	11	*	0	*	4	*	0	*	25	*		*	61	*		*
3	10	*	257	*	11	*	0	*	3	*	0	*	26	*		*	86	*		*
Multi-level		*	84	*	15	*	0	*	7	*	0	*		*		*		*		*
K-3 sum	38	*	1037	*	57	*	0	*	27	*	0	*								
K-3 class average			27	*	1.5	*	0	*	0.7	*	0	*								
K-3 combined child/adult average													101	*		*	249	*		*

*Information unavailable

Table F-5

Toccoa, Georgia

PDC		Comparison	
Head Start Centers	Elementary Schools	Head Start Centers	Elementary Schools
Stephens County →	Toccoa	None	None

All Centers and Schools Combined, by Treatment

	Total classes		Total children		Total teachers				Total paid aides				Child:tchr ratio				Child:aide ratio			
					fulltime		parttime		fulltime		parttime		fulltime		parttime		fulltime		parttime	
	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com
HS	3		47		3				3		1		16				16		47	
K	4		72		2				2				36				36			
1	6		128		6				1		1		21				128		128	
2	5		122		5				1		1		24						122	
3	5		121		5								24							
Multi-level	1		11		1								11							
K-3 sum	24		501		22				7		3									
K-3 class average			21		0.9				0.3		0.1									
K-3 combined child/adult average													23				72		167	

Table F-6

Des Moines, Iowa

PDC		Comparison	
Head Start Centers	Elementary Schools	Head Start Centers	Elementary Schools
Moulton →	Moulton	Elmwood →	Elmwood
		Lucas →	Lucas

All Centers and Schools Combined, by Treatment

	Total classes		Total children		Total teachers				Total paid aides				Child:tchr ratio				Child:aide ratio			
					full-time		part-time		full-time		part-time		full-time		part-time		full-time		part-time	
	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com
HS	2	3	42	41	1	1		1	2	1		1	42	41		41	21	41		41
K	4	4	86	100	2	2			1				43	50			86			
1	4	4	91	86	4	4			1				22.7	21.5			91			
2	3	3.5	65	70	3	3.5			1				21.6	20			65			
3	3	2.5	58	63	3	2.5			1				19.3	25.2			58			
Multi-level																				
K-3 sum	14	14	300	319	12	12		1	4											
K-3 class average			75	79.7	0.8	0.8		0.25	0.2											
K-3 combined child/adult average													25	26.5			75			

Table F-7

Takoma Park, Maryland

PDC		Comparison	
Head Start Centers	Elementary Schools	Head Start Centers	Elementary Schools
Takoma Park →	Takoma Park	New Hampshire Estates →	New Hampshire Estates
		Rolling Terrace →	Rolling Terrace

All Centers and Schools Combined, by Treatment

	Total classes		Total children		Total teachers				Total paid aides				Child:tchr ratio				Child:aide ratio			
					fulltime		parttime		fulltime		parttime		fulltime		parttime		fulltime		parttime	
	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com
HS	4	3	66	57	2	1		1	1	2	2	1	16.5	19		57	66	28.5	33	57
K	6	5	138	52	3	2			2		1		23	26			69		138	
1	4	4.5	113	102	4	4			4				28.2	22.6			28.2			
2	3	2	84	43	3	2			1		3		84	21.5			84		28	
3	2	2	67	59	2	2			1				33.5	29.5			67			
Multi-level	1	4	59	128	2	5					1		59	32					59	
K-3 sum	16	17.5	461	384	14	15			8		5									
K-3 class average			28.8	21.9	0.8	0.8			0.5		0.3									
K-3 combined child/adult average													32.9	25.6			57.6		92.2	

Table F-8

Pontiac, Michigan

PDC			Comparison	
Head Start Centers		Elementary Schools	Head Start Centers	Elementary Schools
Whitmer Human Resource Center	→	Whitmer Human Resource Center	Franklin →	Franklin Frost Longfellow Whittier

	Total classes		Total children		Total teachers				Total paid aides				Child:tchr ratio				Child:aide ratio			
					full-time		part-time		full-time		part-time		full-time		part-time		full-time		part-time	
	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com
HS	4	4	74	73	2	2				4	4		37	36.5			18.5			
K	3	11	70	253	1	5	1	1	1	3			70	50.6	70	253	70	84.3		
1		16		431		17				5				25.4				86.2		
2		14.5		395		14.5								27.2						
3		14.5		370		14.5								25.5						
Multi-level	3	1	211	20	7.3	1			5		1		29	20			42.2		211	
K-3 sum	6	57	281	1469	8.3	52	1	1	6	8	1									
K-3 class average			46.8	253.8	1.4	0.9	0.2	0.02	1	0.2	0.02									
K-3 combined child/adult average													33.8	28.2	281	1469	46.8	183.6	281	

*There are four additional full-time aides who assist the total teaching staff. These aides are included in this ratio.

Table F-9

Del Rio, Texas

PDC		Comparison	
Head Start Centers	Elementary Schools	Head Start Centers	Elementary Schools
Memorial →	Memorial	Lamar →	Lamar

All Centers and Schools Combined, by Treatment

	Total classes		Total children		Total teachers full-time		Total teachers part-time		Total paid aides full-time		Total paid aides part-time		Child:tchr ratio full-time		Child:tchr ratio part-time		Child:aide ratio full-time		Child:aide ratio part-time	
	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com
HS		3		78		3				3				26				26		
K		3		85		3				3				28				28		
1		2		54		2				2				27				27		
2		2		58		2				2				29				29		
3		2		60		2				2				30				30		
Multi-level	12		300		12				10				25				30			
K-3 sum	12	12	300	335	12	12			10	12										
K-3 class average			25	28	1	1			0.8	1										
K-3 combined child/adult average													25	28				28		

Table F-10

Salt Lake City, Utah

PDC		Comparison	
Head Start Centers	**Elementary Schools**	**Head Start Centers**	**Elementary Schools**
	Edison		Emerson
Matheson →	Glendale Park	Matheson →	Hawthorne
	Parkview		Whittier

All Centers and Schools Combined, by Treatment

	Total classes		Total children		Total teachers				Total paid aides				Child:tchr ratio				Child:aide ratio			
					full-time		part-time		full-time		part-time		full-time		part-time		full-time		part-time	
	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com
HS	9	5	170	97	9	5			9	5	6		18.8	19.4			18.8	19.4	28.3	
K	5	8	214	319	4	7	2	2	3	5	4		53.5	45.5	107	159.5	71.3	63.8	53.5	
1	9	12	215	272	9	12			5		3	4	23.8	22.6			43		90.6	68
2	6	9	142	240	6	9			2		4	3	23.6	26.7			71		60	80
3	6	9	168	253	6	9						3	28	28.1						84.3
Multi-level	6	1	149	22	7	1						1	21.3	22						22
K-3 sum	41	39	888	1106	32	38	2	2	10	5	11	11								
K-3 class average			21.7	28.4	0.8	0.9	0.04	0.05	0.24	0.13	0.28	0.28								
K-3 combined child/adult average													27.8	29.1	444	553	88.8	221.2	80.7	100.

Table F-11

Tacoma, Washington

PDC Head Start Centers	PDC Elementary Schools	Comparison Head Start Centers	Comparison Elementary Schools
Lister	→ Lister	Rogers	→ Rogers
		Roosevelt	→ Roosevelt
		Willard	→ Willard

All Centers and Schools Combined, by Treatment

	Total classes		Total children		Total teachers full-time		Total teachers part-time		Total paid aides full-time		Total paid aides part-time		Child:tchr ratio full-time		Child:tchr ratio part-time		Child:aide ratio full-time		Child:aide ratio part-time	
	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com
HS	3	5	60	100	3	2			3	4			20	50			20	25		
K	3	5	94	100	3	5			3	5			31.3	20			31.3	20		
1	3	5	71	100	3	5			3	5			23.6	20			23.6	20		
2	3	5	81	100	3	5			3	5			27	20			27	20		
3	3	5	70	100	3	5			3	5			23.3	20			23.3	20		
Multi-level																				
K-3 sum	12	20	316	400	12	20			12	20										
K-3 class average			23.3	20	1	1			1	1										
K-3 combined child/adult average													26.3	20			26.3	20		

Table F-12

Morgantown, West Virginia

PDC			Comparison		
Head Start Centers		Elementary Schools	Head Start Centers		Elementary Schools
Jerome Park	→	Jerome Park	Cass	→	Cass
Woodburn	→	Woodburn	National	→	National

All Centers and Schools Combined, by Treatment

	Total classes		Total children		Total teachers				Total paid aides				Child:tchr ratio				Child:aide ratio			
					full-time		part-time		full-time		part-time		full-time		part-time		full-time		part-time	
	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com	PDC	Com
HS	2	2	45	40	2	2			2	2	2	2	22.5	20			22.5	20	22.5	20
K	4	4	108	85	2	4			2	4			54				54			
1	2	3	47	69	2	3			1.6	3			23.5	23			29.4	23		
2	1	2	39	60	1	2			.6	2			39	36			65	36		
3	1	2	27	53	1	2			.6	3			27	23.3			45	23.3		
Multi-level	1	1	28	29	1	1			1	1			28				28			
K-3 sum	9	12	249	296	7	12			5.8	13										
K-3 class average			27.7	24.6	0.8	1			.64	1										
K-3 combined child/adult average													35.5	22.4			42.9	22.7		

APPENDIX G

Attrition Findings for Each Site

Attrition Findings for Each Site

California

Before 1974-75 children from the PDC Head Start center advanced to several different elementary schools, and thus the data collected this spring did not permit estimation of attrition among children enrolled in classes from these earlier years. Attrition among children from the 1974-75 PDC Head Start class (now in kindergarten) appears to be 73%. This figure seems so startingly high (and so much at odds with the figures reported for other years) that it was judged advisable to declare the data insufficient for any conclusion. Figures for one of the two comparison centers were available for all four year-groups, beginning in 1971-72, and it is these figures that provide the main basis for estimation of attrition in the comparison centers.

Colorado

No attrition data from Head Start classes were collected this year for several reasons, among which was that no records are kept at the individual elementary schools to indicate whether or not a child attended Head Start. This represents a conscious effort by the schools not to label children. One factor which may affect significantly Pueblo's attrition is that the children are not restricted to Head Start centers or schools by geographic location. Head Start parents are free to select the school their child will attend based on the particular program the school is offering, e.g., IGE, Bilingual, etc. It would be best to encourage parents in the PDC and comparison school geographic locations to enroll their children in those schools.

Connecticut

The data supplied by this site were complete and well-documented. If the indicated trend holds, the expected enrollment of 60 should be adequate for the PDC group and more than adequate for the comparison group. It may be that the attrition data gathered last year showed higher attrition rates because the methods of estimation used by site staff at that time were less precise.

Florida

The data provided by the Florida site appear complete and show a consistently low rate of attrition.

Georgia

The figures obtained for the years prior to 1975-76 reflect a situation which no longer exists, in that children from the county surrounding Toccoa attended Head Start in Toccoa and then returned to their local elementary school. Attendance at the Toccoa Head Start is now limited to Toccoa residents.

Given the change in attendance patterns, retention rates can be expected to be higher than in the past.

Iowa

The attrition data collected appear to be accurate. Although there was no Head Start at Moulton prior to fall 1975, only those children from the Moulton district who attended the Head Start that was operating were included in the attrition calculations.

The comparison school statistics are based on Lucas Elementary only. Many children not in the Lucas district attend the Head Start at Lucas, which reduces the potential sample size, but the situation is even worse at Elmwood, where only 4 children from that district attended the Head Start, which began in 1974-75. Of those 4 children, out of a total of 26, only two remain, and were, therefore, not included in the attrition calculations. Additional comparison elementary schools are to be designated by Iowa in the coming year to increase the number of comparison Head Start children that can be followed longitudinally, but it would still be helpful for the Head Start centers to recruit as many children as possible from the attendance area of the elementary school in which the Head Start is located.

Maryland

The attrition figures for Head Start to third grade appear to coincide closely with the attrition figures from the previous year. No future boundary changes in attendance areas are anticipated, and there have been none since 1971, although some schools have closed.

An attempt should be made to determine the reasons for the extraordinarily high attrition rate in the comparison schools, so that some remedy might be sought.

Michigan

It is uncertain just how much confidence can be placed in the attrition figures from Pontiac, since they are based on incomplete school records. Furthermore, attendance at Pontiac schools has been disrupted by the effects of busing, which has thrown the community into turmoil and increased transiency in the population. However, attrition figures from 1975 and 1976 are in fairly close agreement, and projections are close to the actual figures, indicating that these estimates may have some substance.

Texas

Estimates for this site are probably low in reliability since there was no preschool program prior to 1973-74 (the year of court-ordered integration) and since local records make no distinction between Head Start and non-Head Start programs. Also, there was no Head Start at Memorial (PDC) before this year, so that PDC schools and comparison schools were combined. The 1974-75 records for two schools were missing, introducing greater uncertainty into attrition estimates.

Utah

The attrition figures from Utah appear to be accurate and complete.

Washington

The collection of attrition information appeared thorough and accurate, although the Project Coordinator expressed doubt that attrition is actually as high as the figures made it appear to be.

West Virginia

Attrition information was provided for only one of the two PDC schools, and only for last year, 1974-75, although the Project Coordinator did supply attrition data on low income children in general, which indicated that after they reach kindergarten, attrition is negligible.

Table G-1

Attrition Rate Information: percentage drop estimated over all centers and schools

California

1975 Data	% drop in Head Start from Sept. to Feb.	% drop from K to 1st	%-drop from K to 2nd	% drop from K to 3rd
PDC	*	9	18	11
Comparison	*	2	14	20

1976 Data	% drop from Head Start (74-75) to K	% drop from Head Start (73-74) to 1st	% -drop from Head Start (72-73) to 2nd	% drop from Head Start (71-72) to 3rd
PDC	73			
Comparison	35	26	50	79

*Not available

Table G-2

Attrition Rate Information:
percentage drop estimated over all centers and schools
Colorado

1975 Data	% drop in Head Start from Sept. to Feb.	% drop from K to 1st	% drop from K to 2nd	% drop from K to 3rd
PDC	15	9	41	39
Comparison	15	35	41	40

(No 1976 data available on attrition beginning at Head Start)

Table G-3

Attrition Rate Information: percentage drop estimated over all centers and schools

Connecticut

1975 Data	% drop in Head Start from Sept. to Feb.	% drop from K to 1st	% drop from K to 2nd	% drop from K to 3rd
PDC	6	30	45	62
Comparison	4	25	35	50

1976 Data	% drop from Head Start (74-75) to K	% drop from Head Start (73-74) to 1st	% drop from Head Start (72-73) to 2nd	% drop from Head Start (71-72) to 3rd
PDC	13	14	26	42
Comparison	4	24	17	21

Table G-4

Attrition Rate Information
percentage drop estimated over all centers and schools
Florida

1975 Data	% drop in Head Start from Sept. to Feb.	% drop from K to 1st	% drop from K to 2nd	% drop from K to 3rd
PDC	20	35	25	20
Comparison	10	15	*	*

1976 Data	% drop from Head Start (74-75) to K	% drop from Head Start (73-74) to 1st	% drop from Head Start (72-73) to 2nd	% drop from Head Start (71-72) to 3rd
PDC	15	24	09	*
Comparison	07	04	16	0

*Not available

Table G-5

Attrition Rate Information
percentage drop estimated over all centers and schools
Georgia

1975 Data	% drop in Head Start from Sept. to Feb.	% drop from K to 1st	% drop from K to 2nd	% drop from K to 3rd
PDC	11	30	36	38
Comparison	*	*	*	*

1976 Data	% drop from Head Start (74-75) to K	% drop from Head Start (73-74) to 1st	% drop from Head Start (72-73) to 2nd	% drop from Head Start (71-72) to 3rd
PDC	41	40	40	41
Comparison	*	*	*	*

*Not available

Table G-6

Attrition Rate Information: percentage drop estimated over all centers and schools

Iowa

1975 Data	% drop in Head Start from Sept. to Feb.	% drop from K to 1st	% drop from K to 2nd	% drop from K to 3rd
PDC	20	11	36	48
Comparison	4	19	44	61

1976 Data	% drop from Head Start (74-75) to K	% drop from Head Start (73-74) to 1st	% drop from Head Start (72-73) to 2nd	% drop from Head Start (71-72) to 3rd
PDC	21	50	68	74
Comparison	50	45	60	55

Table G-7

Attrition Rate Information:
percentage drop estimated over all centers and schools
Maryland

1975 Data	% drop in Head Start from Sept. to Feb.	% drop from K to 1st	% drop from K to 2nd	% drop from K to 3rd
PDC	13	49	65	64
Comparison	5	21	67	72

1976 Data	% drop from Head Start (74-75) to K	% drop from Head Start (73-74) to 1st	% drop from Head Start (72-73) to 2nd	% drop from Head Start (71-72) to 3rd
PDC	31	67	65	62
Comparison	38	51	59	93

Table G-8

Attrition Rate Information: percentage drop estimated over all centers and schools

Michigan

1975 Data	% drop in Head Start from Sept. to Feb.	% drop from K to 1st	% drop from K to 2nd	% drop from K to 3rd
PDC	8	27	46	50
Comparison	1	*	*	*

1976 Data	% drop from Head Start (74-75) to K	% drop from Head Start (73-74) to 1st	% drop from Head Start (72-73) to 2nd	% drop from Head Start (71-72) to 3rd
PDC	66	53	56	65
Comparison	16	31	15	32

*Data for comparison schools was not submitted.

Table G-9

Attrition Rate Information:
percentage drop estimated over all centers and schools
Texas

1975 Data	% drop in Head Start from Sept. to Feb.	% drop from K to 1st	% drop from K to 2nd	% drop from K to 3rd
PDC	*	*	*	*
Comparison	*	*	*	*

1976 Data	% drop from Head Start (74-75) to K	% drop from Head Start (73-74) to 1st	% drop from Head Start (72-73) to 2nd	% drop from Head Start (71-72) to 3rd
PDC	31**	30**	*	*
Comparison			*	*

*No preschool program prior to 1973-74.

**Local records make no distinction between Head Start and non-Head Start programs.

Table G-10

Attrition Rate Information: percentage drop estimated over all centers and schools

Utah

1975 Data	% drop in Head Start from Sept. to Feb.	% drop from K to 1st	% drop from K to 2nd	% drop from K to 3rd
PDC	23	25	34	45
Comparison	21	24	46	50

1976 Data	% drop from Head Start (74-75) to K	% drop from Head Start (73-74) to 1st	% drop from Head Start (72-73) to 2nd	% drop from Head Start (71-72) to 3rd
PDC	35	57	62	65
Comparison	56	83	81	89

Table G-11

Attrition Rate Information: percentage drop estimated over all centers and schools

Washington

1975 Data	% drop in Head Start from Sept. to Feb.	% drop from K to 1st	% drop from K to 2nd	% drop from K to 3rd
PDC	42	33	50	65
Comparison	5	50	50	65

1976 Data	% drop from Head Start (74-75) to K	% drop from Head Start (73-74) to 1st	% drop from Head Start (72-73) to 2nd	% drop from Head Start (71-72) to 3rd
PDC	46	67	69	80
Comparison	56	76	88	80

Table G-12

Attrition Rate Information: percentage drop estimated over all centers and schools

West Virginia

1975 Data	% drop in Head Start from Sept. to Feb.	% drop from K to 1st	% drop from K to 2nd	% drop from K to 3rd
PDC	10	24	28	20
Comparison	10	13	13	10

1976 Data	% drop from Head Start (74-75) to K	% drop from Head Start (73-74) to 1st	% drop from Head Start (72-73) to 2nd	% drop from Head Start (71-72) to 3rd
PDC	39	**	*	*
Comparison	17	34	33	

*No Head Start or Kindgergaten

**No Head Start

APPENDIX H

COMPOSITION OF SUBSCALES FOR THE PUPIL OBSERVATION CHECKLIST (POCL) AND THE PDC CHILD RATING SCALE (CRS), BASED ON FACTOR ANALYSES

Subscales of the POCL

Factor analysis of the Pupil Observation Checklist resulted in two factors, which have been named "Task Orientation" and "Extroversion" to describe the common characteristics of the item loading on the factors. The items loading on each of these factors and the factor loadings are shown in Table H-1. Scores on the subscales were calculated by summing the actual scores on all items for a subscale. The subscale scores were then used in all subsequent analyses of POCL spring data.

Subscales of the PDC Child Rating Scale

The items of the PDC Child Rating Scale were factor analyzed using the same method as for the POCL. The five resulting factors have been assigned the names, "Poise," "Aggressiveness," "Determination," "Self-Sufficiency," and "Flexibility." Table H-2 shows the factor on which each item loaded highest and the item's loading for that factor. As with the POCL, subscale scores for use in subsequent analyses were constructed by summing actual scores on all items for a subscale, using negative scores where loadings were negative.

Table H-1

Subscales of the POCL, Based on Factor Analysis[1] of the Items

Spring 1976

POCL Item	Loading of Item on Each Factor; Highest Loading in Italics	
	Factor 1 Task Orientation	Factor 2 Extroversion
Cooperative	*.79*	.29
Sociable	.37	*.79*
Involved	*.73*	.40
Talkative	.18	*.83*
Attentive	*.77*	.20
Active	.19	*.71*
Relaxed	*.60*	.39
Quick to Respond	*.82*	.31
Attempts Difficult Tasks	*.83*	.17
Keeps Trying	*.86*	.19
Realistically Self-Confident	*.72*	.22

---

[1] Principal Components factor analysis, varimax rotation.

Table H-2

Subscales of the PDC Child Rating Scale, Based on Factor Analysis[1], Spring 1976 Data

Compared to other children of the same age and background, how often does this child behave in the following ways?	Factor for Which Loading Was Highest					
	5. Flexibility	4. Self-Sufficiency	3. Determination	2. Aggressiveness	1. Poise	Highest Loading
1. Uses words or wits to try to influence others	5	4	3	2	[✓ 1]	.45
2. Uses physical force to try to control others	5	4	3	[✓ 2]	1	.79
3. Succeeds in influencing or controlling others	5	4	3	2	[✓ 1]	.59
4. Is controlled or influenced by others	5	4	[✓ 3]	2	1	-.54
5. Imitates others or follows them around	5	4	[✓ 3]	2	1	-.66
6. Competes with others for toys, attention, achievement	5	4	3	[✓ 2]	1	..55
7. Asserts his/her rights to fair treatment	5	4	3	[✓ 2]	1	.40
8. Accepts or abides by school, classroom rules	5	4	3	[✓ 2]	1	-.68
9. Cooperates and shares with others	5	4	3	[✓ 2]	1	-.65
10. Shows verbal affection to others	5	4	3	2	[✓ 1]	.65
11. Shows verbal dislike or hostility to others	5	4	3	[✓ 2]	1	.70
12. Shows physical affection to others	5	4	[✓ 3]	2	1	-.50
13. Shows physical dislike or hostility to others	5	4	3	[✓ 2]	1	.79
14. Shows self-confidence	5	4	3	2	[✓ 1]	67
15. Shows awareness of and pride in own culture	5	4	3	2	[✓ 1]	.64
16. Decides for self what to do, with whom to play	5	[✓ 4]	3	2	1	.57
17. Gets the attention of adults appropriately	5	[✓ 4]	3	2	1	.44
18. Gets the attention of peers appropriately	5	4	3	2	[✓ 1]	.47

---

[1] Principal Components factor analysis, varimax rotation

Table H-2
(continued)

Compared to other children of the same age and background, how often does this child behave in the following ways?	Factor for Which Loading Was Highest					
	5. Flexi-bility	4. Self-Sufficiency	3. Deter-mination	2. Aggres-siveness	1. Poise	Highest Loading
19. Attempts to solve social problems with little adult assistance	5	4	3	2	[✓ 1]	.54
20. Takes on the role of adult during dramatic play	5	4	3	2	[✓ 1]	.55
21. Shows respect for or tolerance of others' ideas and behavior or looks . . .	[✓ 5]	4	3	2	1	.46
22. Recognizes others' feelings, responds appropriately	5	4	3	2	[✓ 1]	.52
23. Talks freely to children.	5	4	3	2	[✓ 1]	.77
24. Talks freely to adults.	5	4	3	2	[✓ 1]	.70

Table H-2
(continued)

Factors for Which Loading Was Highest

Compared to other children of the same age and background, how often does this child behave in the following ways?	5. Flexibility	4. Self-Sufficiency	3. Determination	2. Aggressiveness	1. Poise	Highest Loading
25. Works or plays well on his/her own	5	[✓ 4]	3	2	1	.57
26. Uses alternative strategies if initial problem-solving methods fail	[✓ 5]	4	3	2	1	.61
27. Is easily distracted when doing a task . .	5	4	[✓ 3]	2	1	-.53
28. Asks for needed help to do a task.	5	4	[✓ 3]	2	1	-.53
29. Asks for unnecessary help to do a task . .	5	4	[✓ 3]	2	1	-.70
30. Sees his/her errors in a task and corrects them.	[✓ 5]	4	3	2	1	.71
31. Gives up on tasks before they are finished	5	4	[✓ 3]	2	1	-.61
32. Returns to unfinished tasks after interruption	[✓ 5]	4	3	2	1	.54
33. Enjoys tasks he/she chooses.	5	[✓ 4]	3	2	1	.82
34. Shows pride in what he/she does or makes.	5	[✓ 4]	3	2	1	.75
35. Chooses tasks that are too easy for him/her.	5	4	[✓ 3]	2	1	-.56
36. Chooses tasks that are too hard for him/her.	5	4	3	2	1	--
37. Expects to succeed, not afraid of failure.	[✓ 5]	4	3	2	1	.58
38. Has a desire to master all kinds of skills	5	[✓ 4]	3	2	1	.56
39. Needs much encouragement or material rewards to attempt tasks	5	4	[✓ 3]	2	1	-.59